UNDERSTANDING THE STRUCTURAL BASIS FOR FUNCTIONAL DIFFERENCES IN STAPHYLOCOCCAL MSCRAMMS SDRE1 AND BBP/SDRE2 AND THEIR ROLE IN SPECIES TROPISM

A Dissertation

by

MATHEW PRASHANTH FRANCIS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Magnus Höök
Committee Members,	Sarah Bondos
	Burton Dickey
	David Huston
	Yi Xu
Head of Department,	Van Wilson

May 2015

Major Subject: Medical Sciences

ABSTRACT

Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMMs) on *Staphylococcus aureus* play roles in attachment, invasion and immune evasion. It has been previously reported that bone sialoprotein-binding protein (Bbp) binds to human fibrinogen. Herein, we show that Bbp and SD-repeat protein E (SdrE) are allelic variants whose *in vitro* fibrinogen binding profile provides a rationale for the epidemiological presence of Bbp and SdrE in human and animal Staphylococcal strains. Epidemiological studies show that *bbp* is found in 32% of human staphylococcal strains, whereas it is nearly absent from animal staphylococcal strains. We show through basic and advanced *in vitro* biochemical techniques that Bbp shows a vastly higher affinity for human fibrinogen than SdrE and that Bbp specifically displays affinity for fibrinogen from humans. In contrast, SdrE shows varying affinity for fibrinogen from a wide array of various animals.

The structural basis for this difference is elucidated by determining the molecular structures of Bbp and SdrE with no ligands as well as the molecular structures of Bbp and SdrE in complex with a peptide representing the fibrinogen ligand binding sequence. Data from these structural analyses were used to inform mutational analysis of the recombinant proteins, both to confirm the integrity structural models and to discover which residues are responsible for

the differences in binding profile. From these data, we show that the Lock Regions of the two proteins are a major cause of this difference in binding profile and species tropism.

ACKNOWLEDGMENTS

This research was carried out at the Institute of Biosciences and Technology (IBT) under the guidance of Dr. Magnus Höök from 2009 – 2013. Support for the research was provided by grant R01AI020624 from the National Institute of Allergy and Infectious Disease under Dr. Magnus Höök. Dr. Van Wilson and Dr. Leibowitz also provided part of my stipend from the Texas A&M Health Science Center graduate student fund.

First, I would like to thank my mentor, Dr. Magnus Höök, for allowing me to join his lab and work under his mentorship over the past four years. I became interested in his research because it represented a chance to combine basic science biochemical research with clinically relevant topics. During my time in his lab, Dr. Höök has had an open door policy that allowed for conversations about everything from theoretical biochemistry to work-life balance to the changing funding environment. He has been an excellent mentor for an MD/PhD student who has different goals, requirements and hard deadlines that do not always align with the more normal graduate school path. I am very grateful for the flexibility that he allowed me in scheduling and pursuing topics of interest. I would also like to thank the members of my advisory committee, Dr. Sarah Bondos, Dr. Burton Dickey, Dr. David Huston, and Dr. Yi Xu, for their guidance, suggestions and evaluations of my projects. I am grateful to have the

opportunity to have an advisory committee whose expertise ranges from basic science research to clinical research.

I am incredibly grateful to the present and past members of the Höök laboratory who have trained me in myriad techniques and devices and proofread countless pages. I would like to thank Dr. Cana Ross, whose guidance through the highs and lows of graduate school were pivotal. I would also like to thank Dr. Ross and Dr. Ana Cohen for their help in cloning, Xiaowen Liang, who performed all SPR experiments and instructed me in the use of many new techniques, and Dr. Ganesh Vannakambadi, who performed the crystallization and structural work. Furthermore, I would like to thank collaborators Dr. Ed Feil at the University of Bath, Dr. Ross Fitzgerald at the University of Edinburgh, and Dr. Paal Andersen of the Statens Serum Institut in Denmark. Also, I would like to thank newer additions to our lab in Dr. Brandon Garcia and Dr. Danielle McGrath for their help in editing this thesis and in the lab.

My progress to this point in my career would not be possible without my wife, Lauren Francis, who has supported me in all ways, has maintained confidence in me even when I did not have it myself, and has partnered with me in raising our son, Grayson. I would also like to thank my parents, whose hard work, sacrifice and belief in me can never be repaid.

NOMENCLATURE

Abs_{280}, Abs_{450}: Absorbance at 280 nm or 450 nm

Bbp: BSP-Binding Protein

BSA: Bovine Serum Albumin

BSP: Bone Sialoprotein

CA-MRSA: Community Acquired – MRSA

CD: Circular Dichroism

ClfA: Clumping Factor A

ClfB: Clumping Factor B

DAP: Daptomycin

DAP-R: Daptomycin Resistant

DNA: Deoxyribonucleic Acid

ECM: Extracellular Matrix

ELISA: Enzyme-Linked Immunosorbent Assay

Fg: Fibrinogen

fH : Factor H

Fn: Fibronectin

FnbpA: Fibronectin-binding Protein A

FnbpB: Fibronectin-binding Protein B

HA-MRSA: Hospital Acquired – MRSA

IPTG: Isopropyl β-D-1-thiogalactopyranoside

ITC: Isothermal Titration Calorimetry

IV: Intravenous

LA-MRSA: Livestock Acquired – MRSA

MRSA: Methicillin-Resistant *Staphylococcus aureus*

OE-PCR: Overlap Extension – Polymerase Chain Reaction

PBP2: Penicillin Binding Protein 2

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

RNA: Ribonucleic Acid

SDM: Site Directed Mutagenesis

SEC: Size Exclusion Chromatography

SPR: Surface Plasmon Resonance (Biacore)

TSS: Toxic Shock Syndrome

TSST-1: Toxic Shock Syndrome Toxin-1

VISA: Vancomycin-Insensitive *Staphylococcus aureus*

VRSA: Vancomycin-Resistant *Staphylococcus aureus*

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

STAPHYLOCOCCUS AUREUS

Staphylococcus aureus is a non-motile, Gram-positive bacterium that colonizes humans and causes a wide array of pathology. It was first isolated by surgeon Sir Alexander Ogston from the pus of surgical abscesses. Ogston used eggs to isolate cultures of the bacteria and then recreated the abscess phenotype in rabbits. When he reported his findings in 1880, he named the bacteria Staphylococcus.[1] It derives its name from its microscopic appearance as a cluster of grapes and its macroscopic golden coloring. As with other Gram-positive bacteria, *S. aureus* has a thick cell wall composed of peptidoglycan and lipoteichoic acid. *S. aureus* is differentiated from other Gram-positive bacteria such as *Streptococci* and *Enterococci* by the presence of the enzymes catalase and coagulase. While *S. aureus* is an aerobic bacterium that grows best in temperatures found in the human body, 30-37 °C, it is capable of surviving in a variety of environments including conditions of low oxygen, high osmotic pressure, and broad temperatures.[2]

Staphylococcus aureus in human pathogenesis

At any given moment, 30-50% of the US population is colonized with *S. aureus*. 20-30% of the population are consistent carriers, while the remaining subset is composed of intermittent carriers. Carriers display increased risk for diseases caused by the bacterium. *S. aureus* most frequently colonizes the anterior nares, skin and nails.[2] It is an opportunistic pathogen; once the outer or initial host defenses are compromised, a diverse set of pathologies can result. These pathologies range from generally benign skin infections such as boils and furuncles to potentially lethal conditions including necrotizing fasciitis and pneumonia. Patients with comorbidities such as diabetes, acquired immunodeficiency syndrome (AIDS), or recent invasive medical procedures are at still greater risk of infection.[2]

S. aureus infections are frequently grouped by their source location: healthcare-associated MRSA (HA-MRSA), community-associated MRSA (CA-MRSA), and the newer finding of healthcare-associated with a community onset MRSA. Historically, MRSA infections have largely been attributed to HA-MRSA, but the incidence and proportion of CA-MRSA has been steadily increasing over the past decade.[3] In North America and parts of Latin America, this can be correlated directly with the rise of the USA300 strain, the spread of which has reached pandemic proportions.[4] The virulence of the USA300 strain resulted in MRSA being the leading cause of death in the United States in 2005 by an infectious agent, surpassing HIV/AIDS.[5] A recent epidemiological study of

MRSA in Canada found that CA-MRSA increased from 19.7% of MRSA infections in 2007 to 36.4% of infections in 2011.[6]

S. aureus displays the remarkable ability to cause disease in broadly different organs and tissues. One of the most common types of infection caused by *S. aureus* is skin and soft tissue infections (SSTIs), which include folliculitis, furunculosis and abscesses. The growing incidence of MRSA SSTIs is correlated to the spread of CA-MRSA and has resulted in an increase in number of patient visits to emergency departments with SSTIs; severe SSTIs such as necrotizing fasciitis require both immediate treatment with the broad spectrum antibiotics that include MRSA coverage.[7] While broad spectrum antibiotics can be administered over short periods of time to manage emergent cases, antibiotic selection for the long term management of recurrent SSTIs is difficult due to the need to balance concerns about antibiotic resistance, long term side effects, and efficacy. Consistent decolonization is the goal, but clearance of the bacteria is frequently only temporary.

S. aureus is also a major causative agent of osteomyelitis across all age groups and comorbidities. While osteomyelitis infections contain multiple pathogens a majority of the time, *Staphylococcus aureus* is the leading cause and is found in 50-70% of osteomyelitis cultures. The rise in methicillin-resistant *S. aureus* (MRSA) incidence has been cited as the biggest epidemiologic challenge in osteomyelitis and osteoarticular infections. The route of infection varies by age.[8] In pediatric patients, secondary spread via a hematogenous route

is the most common route. However, in adults the most common routes are primary infection due to trauma or surgical procedures or secondary infection from local spread from a SSTI. One notable exception is the seeding of osteomyelitis of the spine in adults, which most frequently occurs via hematogenous route.

Chronic or recurring osteomyelitis is another challenge posed by MRSA infection of bone tissue or orthopedic implants. In these settings, the generation of a biofilm is often the underlying cause of reduced bacterial clearance. Biofilm forms via the deposition of an extracellular matrix composed of proteins, carbohydrates and nucleic acids, creating a protected environment for the bacteria. Biofilms also pose grave challenges in treatment due to significantly higher resistance to antibiotics; it has been estimated that bacteria in a biofilm require levels of antibiotics an order of magnitude higher or greater to be effective. In the case of prosthetic device or orthopedic implant, the synthetic material must often be removed.[9]

MRSA pneumonia has represented a serious health threat for many years. HA-MRSA presents one of the more difficult challenges in infectious disease due to the difficulty in adequately sterilizing hospital equipment such as mechanical ventilators. By 2002, a new trend was observed of MRSA pneumonia increasingly caused by CA-MRSA strains instead of solely being the realm of HA-MRSA. Specifically, the rapid spread of the USA300 strain in the community was altering the course of the typical infection. CA-MRSA is now recognized as

causing a rapidly progressing, frequently fatal necrotizing pneumonia. It is characterized by diffuse inflammation, vascular leakage, and severe tissue destruction in the host lungs. Unfortunately, epidemiological studies of the spread of CA-MRSA pneumonia are difficult to significant variations across regions.[10]

In addition to causing an array of pathologies by dissemination and seeding of different organs, *S. aureus* is capable of pathogenesis via exotoxin production. *S. aureus* is the second leading cause of gastroenteritis. *S. aureus* gastroenteritis is classified according to the presence of 7 known enterotoxins (staphylococcal Exotoxins A-H).[11] This exotoxin-based disease is notable for its rapid onset, ranging from 0.5 to 8 hours after ingestion of the contaminated source, as well as nausea, cramps and diarrhea. While there are few deaths in wealthy nations from this disease, the prevalence of the disease warrants further research and monitoring.[12]

Furthermore, *S. aureus* also causes Toxic Shock Syndrome (TSS) via the pyrogenic superantigen termed Toxic Shock Syndrome Toxin-1 (TSST-1). TSST-1 functions by binding to the α-chain of MHC Class II via a specific orientation that results in the activation of the T-cell receptor despite the absence of its specific ligand. This causes a broad, dramatic, non-specific activation of CD4 and CD8 T-cells which results in a large release of interleukin-1, interleukin-2 and TNFα. The rapid increase of these signals causes a strong acute phase reaction including high fever, low blood pressure, malaise and coma.[13] While Toxic Shock

Syndrome is traditionally associated with improper use of feminine hygiene products, TSS is also growing problem in pediatric burns where it causes unexpected deaths in patients. Recently, pediatric burn TSS have become the most common subtype of this syndrome.[14]

Infective endocarditis is another disease that can be caused by *S. aureus*. Traditionally, *S. aureus* endocarditis is a right sided endocarditis that is associated with IV drug abuse as this provides a direct route of entry from contaminated skin to the heart valve. This form of infective endocarditis has high morbidity and mortality with mortality rates seen in the range of 20-30%.[15] Similar to osteomyelitis and orthopedic implant infection, *S. aureus* can colonize and infect both native valves and prosthetic valves. In cardiovascular implant infection, the coating of implants with serum factors such as fibrinogen and fibronectin allow for targeting by *S. aureus* adhesins. The bacteria can subsequently form a biofilm on the implanted material that results in an infection that is significantly more difficult to treat. In most cases, the implanted heart valve must be removed because it is not possible to safely raise serum levels of antibiotics to the necessary level for clearance.[16]

Multifaceted antibiotic resistance in MRSA

S. aureus infections are frequently difficult to treat clinically due to the wide and variable profile of antibiotics against which the bacteria have gained resistance. *S. aureus* epidemics have occurred in waves directly correlated to the generation and spread of antibiotic resistance mechanisms by the bacteria. The first wave occurred when penicillinase-containing *S. aureus* reached pandemic levels in both hospital and community settings in the 1950s. Methicillin, a semisynthetic molecule that is derived from penicillin but resistant to penicillinase, was developed to combat this wave and did so effectively for a time. However, the next epidemic wave of *S. aureus* occurred less than a year after the introduction of methicillin because of the spread of a new resistance gene - penicillin binding protein 2 (*pbp*2). This methicillin-resistant *S. aureus* epidemic wave was restricted to Europe, but was then replaced in the 1980s by novel MRSA lineages that lead to the global pandemic that is seen today.[17]

Methicillin resistance is now carried on a mobile genetic element called the Staphylococcal Cassette Chromosome *mec* (SCC*mec*). There are a variety of SCC*mec* element types that range 21 to 67 kb, but they all carry *mecA*, a gene that encodes for PBP2 and bestows resistance to methicillin and other similar cell wall inhibitors upon the bacteria. SCC*mec* elements also contain *ccr* genes that code for recombinases which allow for the integration and excision of the mobile genetic element into and out the *S. aureus* genome. The categorization of SCC*mec* into types is performed based on the nucleotide sequences of the *mecA*

and *ccr* genes. Various other genes can be on a particular SCC*mec*, including other virulence factors. Due to this, tracking of the distribution of the different SCC*mec* elements is an important epidemiological goal.

In addition to methicillin and similar antibiotics, MRSA is resistant to erythromycin, clindamycin, ciprofloxacin, tetracycline, and other antibiotics. The current drug that is most effective at treating MRSA is vancomycin, a glycopeptide cell wall inhibitor. This antibiotic is not ideal for treatment because it must be administered via IV due to inefficient gut absorption, must be used at high concentrations, and has potential nephrotoxic side effects. Additionally, there have isolated case reports of vancomycin-intermediate *Staphylococcus aureus* and vancomycin-resistant *Staphylococcus aureus*. One of the mechanisms by which this resistance occurs is by alteration of the pentapeptide sequence is used in cell wall production.

Another antibiotic that had proved effective against some MRSA strains is daptomycin (DAP), which is a lipoprotein that targets the cell membranes of the bacteria. Daptomycin usage is restricted by the fact that it is inhibited by pulmonary surfactant, making it ineffective against MRSA pneumonia. Additionally, numerous recent reports show that mechanisms of daptomycin resistance (DAP-R) are emerging and spreading amongst MRSA isolates.[18] While other options like linezolid, trimethoprim-sulfamethaxazole (Bactrim), and tigecycline are available in certain clinical settings, there is a clear need for the development of the next generation of antibiotics that will target MRSA.

LIVESTOCK ACQUIRED – METHICILLIN RESISTANT *Staphylococcus aureus*

While *Staphylococcus aureus* is an opportunistic pathogen that colonizes humans, there are members of the *Staphylococci* genus whose hosts are non-human species. *Staphylococcus gallinarum* is observed in avian species such as chickens and pheasants, *Staphylococcus hyicus* is found most frequently in bovine species, *Staphylococcus lentus* occurs most frequently in ovine and caprine species, and *Staphylococcus intermedius* and *Staphylococcus pseudintermedius* are most commonly associated with canine species. Despite other members of the *Staphylococcus* family having adapted to non-human species, *Staphylococcus aureus* is able to colonize and cause disease in a wide variety of animals.[19]

The ability of *S. aureus* to spread to other species has become a global health issue. While the ability of MRSA to colonize animals has been described since 1972, the emergence of clonal complex 398 (CC398) has resulted in the widespread epidemic of Livestock Associated–MRSA infections (LA-MRSA) across multiple continents and myriad species. The origins of LA-MRSA CC398 were not clear at first; originally, models that proposed that this represented a new bacterial species that combined traits from multiple species of *Staphylococcus*. However, recent data from whole genome sequencing have been used to create a phylogenetic reconstruction of the LA-MRSA CC398 clonal

lineage. These data show that the origin of this LA-MRSA clonal complex is a methicillin sensitive *S. aureus* strain from humans.[20]

The spread of LA-MRSA has led to two major public health concerns. First, the rapid spread of this clonal complex has resulted in the presence of drug resistant *S. aureus* in a variety of raw meat products across North America and Western Europe. One Dutch study showed that LA-MRSA is present in 12% of 2217 raw meat samples taken from retail stores in the Netherlands with a range of 10-16% varying by animal subset.[21] A similar study performed in Iowa, USA found LA-MRSA in 16.4% 165 raw meat samples taken from retail stores with a 6-19% range depending on animal subset.[22]

The second major concern is LA-MRSA transitioning from livestock species back into humans. The ability of LA-MRSA to colonize the humans who handle livestock was first reported in the early 2000s and has been a growing problem for livestock handlers and their families.[23] There have been isolated cases of LA-MRSA occurring in humans who are not in contact with livestock animals, but there is no evidence to date that LA-MRSA is able to spread via human to human interactions on the larger scale seen with other MRSA sequence types. One recent report noted that human-to-human transmission of a pig-associated clade of LA-MRSA CC398 contained a unique φ7 bacteriophage that is likely example of horizontal transfer between human-associated MRSA CC398 and LA-MRSA CC398.[24]

After *S. aureus* is transferred from humans into a new animal host, *S. aureus* adapts to its new host. There are multiple mechanisms for this adaptation. First, *S. aureus* either accumulates premature stop codons in, or completely loses the genes for, virulence factors that are specific to the human host (Figure 1-1). This is a reflection of the evolutionary pressure that *S. aureus* and other bacteria face to keep their genomes free of ineffective genes. Examples of this include the loss of φSa3, a mobile genetic element in the *S. aureus* accessory genome that contains genes such as *chp* and *scn* that regulate the human immune system. Uhlemann *et al* performed whole genome sequencing on LA-MRSA S3085, an isolate from pigs, and discovered that MSCRAMMs such as *clfA, clfB,* and *fnbpB* were truncated.[25] In another study of avian staphylococcal isolates, it was determined by Lowder *et al* that the initial strain came from humans, made the jump to an avian species approximately 36 years ago, and that *sPA* had accumulated a nonsense mutation.[26]

In addition to losing human specific virulence factors, *S. aureus* bacteria that have made the jump to a new animal host also gain virulence factors that have been shown to be or are predicted to be virulence factors for the new host. An example of this is the presence of φAvβ3 in LA-MRSA CC398 isolates that come from raw turkey meat products, but was found only rarely in LA-MRSA CC398 isolates from other animals. This prophage contains the SAAV_2008 and SAAV_2009 genes that belong to the avian-niche-specific accessory gene pool.[25]

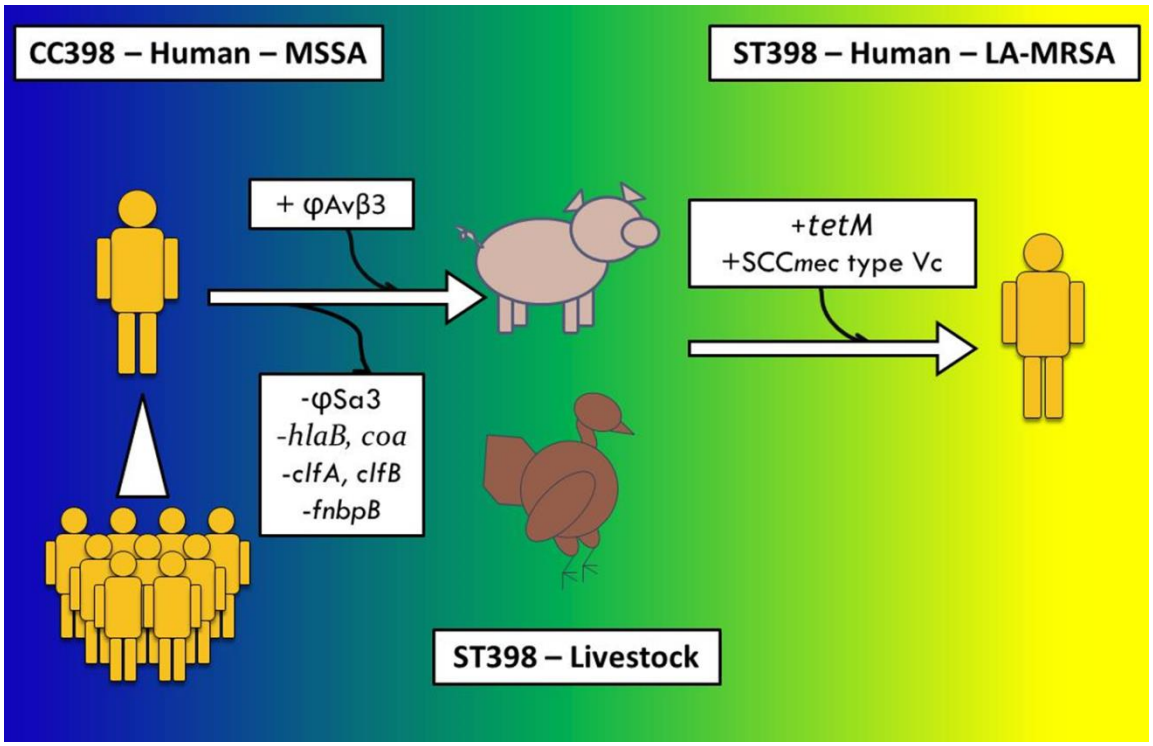


Figure 1-1. Host Adaptation in LA-MRSA CC398.

The underlying mechanism of the acquisition of antibiotic resistance in the ST398 clones from animal hosts is related to livestock production practices, specifically the food and antibiotics that are given to the animals. The original MSSA isolates that made the jump to humans do not show broad antibiotic resistance. However, the administration of tetracycline resulted in the selection of bacteria positive for the tetracycline-resistance gene *tetM*. 99% of the livestock-associated isolates sampled were positive for *tetM*, yet none of the original human-associated isolates contain them. However, the emergence of tetracycline-resistance does not correlate with the emergence of broad multidrug

resistance. Broad multidrug resistance has been shown to be achieved when the livestock animals are given broad-spectrum cephalosporins; these results have spurred the FDA to renew efforts to greatly restrict the use of cephalosporins in livestock. Interestingly, a recent publication suggests that the use of metals such as zinc in animal feed selects for MRSA CC398 strains. The *S. aureus czrC* gene provides resistance to cadmium, zinc and other metals and was recently found to be located on the SCCmec type Vc. This specific SCCmec is present in a majority of LA-MRSA CC398 isolates.[20,27]

There are a growing number of examples of LA-MRSA CC398 infecting livestock handlers. To date, there is little evidence of these bacteria subsequently spreading from these patients to other patients. This suggests that by spending time in an animal host and adapting to that new host environment, *S. aureus* has reduced virulence in humans. This is likely due to the loss or truncation of human specific virulence factors.

MSCRAMMs

Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMMs) are a class of adhesion proteins that are anchored to the surface of Gram-positive bacteria and play a major role in adhesion, as well as other roles including host immune modulation and the disruption of hemostasis. MSCRAMMs are generally expressed in the log phase of growth, reflecting a role

in early adhesion. In contrast, the following stationary phase sees the expression of genes like secreted toxins. After expression, MSCRAMMs are targeted for secretion via the 30 amino acid signal sequence at the amino terminus of the protein. Subsequently, MSCRAMMs are covalently attached to the cell wall by a class of enzymes termed sortases. In *S. aureus*, Sortase A recognizes the LPXTG motif on secreted MSCRAMMs, and then enzymatically cleaves this sequence between the threonine and glycine residues. After the cleavage event, Sortase A attaches the MSCRAMM to the peptidoglycan layer via a transpeptidase reaction. This covalent attachment results in the carboxy terminus of the MSCRAMMs being attached to the cell wall.[28]

With the covalent linkage occurring near the carboxy terminus, the closer a domain of an MSCRAMM is to the amino terminus, the farther that domain will protrude from the surface of the cell. Moving outward from the cell wall anchoring domain, there is a cell wall spanning domain (W) which forms the support architecture necessary to display the ligand binding domains at a farther distance from the surface of the cell. This role is also accomplished by SD repeats (R) and B-repeats in some MSCRAMMs. The ligand binding A domain can frequently be further divided into N1, N2, and N3 domains (Figure 1-2).[29]

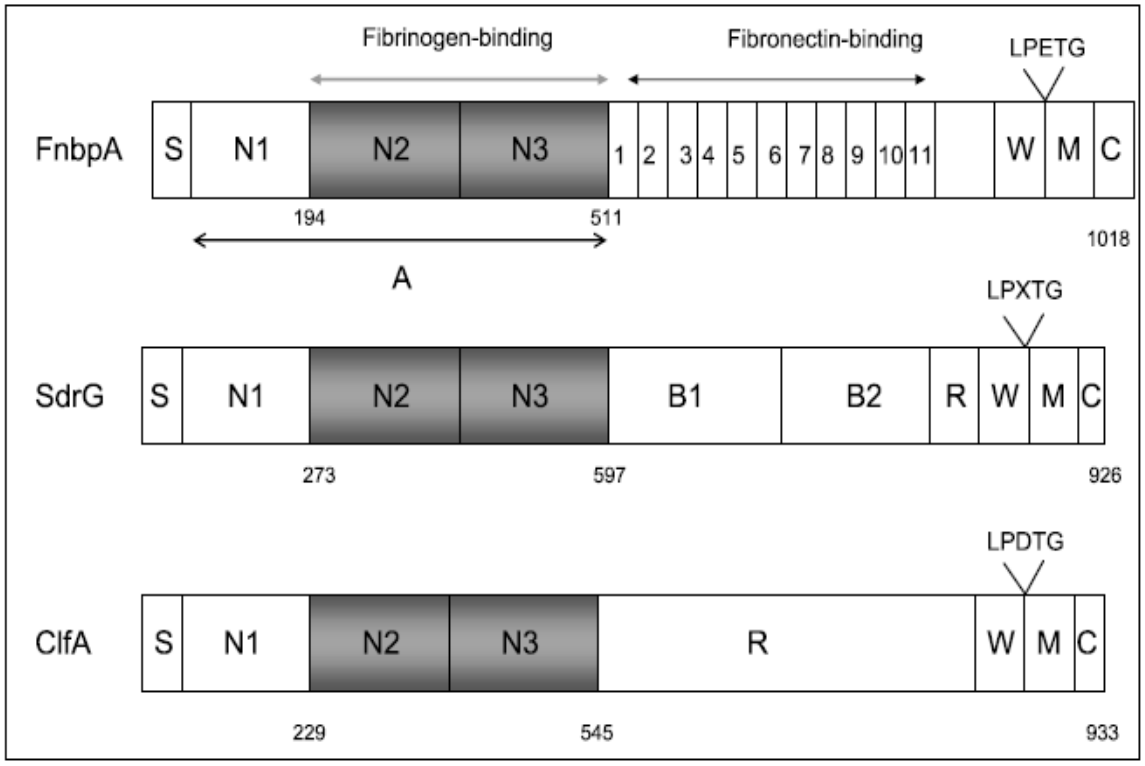


Figure 1-2. Schematic of Fibrinogen-binding MSCRAMMs. (M. Höök)

The presence of the LPTXG motif in MSCRAMMs has allowed for the discovery of 22 of the cell wall anchored proteins. The Höök lab is amongst many labs that have studied these MSCRAMMs to understand their role in attachment and adhesion to host tissues. This work has uncovered numerous host ligands that are targeted by MSCRAMMs, including collagen, fibrinogen, fibronectin, bone sialoprotein and other proteins of the extracellular matrix. There is a surprisingly high level of redundancy in host proteins that are targeted, but the

MSCRAMMs frequently target different regions of the same protein. The evolutionary pressure the bacteria are under to maintain an efficient genome suggests that host ligands that are targeted by more than one MSCRAMM are very important for colonization or virulence.

Interestingly, most MSCRAMMs are capable of binding to multiple host proteins (Table 1-1). While there is a clear incentive for the bacteria to have multifaceted adhesins on its surface, the development of molecules that are able to bind with high affinity via a lock and key method to multiple ligands is remarkable from a structural perspective.

FnbpA and FnbpB

FnbpA and FnbpB are two MSCRAMMs found in *S. aureus* that are named for their ability to bind to the extracellular matrix protein fibronectin type I. This binding has been shown to occur through a tandem β-zipper mechanism. Both FnbpA and FnbpB also have fibrinogen binding domains that are distinct from their fibronectin binding domains; additionally, FnbpA has been shown to bind to elastin via its fibrinogen binding site. FnbpA has been shown to play a role in infective endocarditis and cardiovascular device infections.

Table 1-1. MSCRAMMs and their multiple ligands. (V. Vazquez)

MSCRAMM	Ligands
Protein A	IgG, von Willebrand Factor, TNFα receptor (TNFR), epidermal growth factor receptor (EGFR)
FnbpA	fibronectin, fibrinogen (Fg), elastin
FnbpB	fibronectin, fibrinogen
ClfA	fibrinogen, complement regulator factor I
ClfB	fibrinogen, cytokeratin 10

Over the past decade, the rate of *S. aureus* infection of these devices has grown rapidly. These infections occur because of the ability of *S. aureus* to first bind to the device and then lay down a biofilm. Fibronectin plays a critical role in this initial attachment because it is one of the first host proteins to encounter and coat the implanted device. After initial attachment to the host protein-coated device, the bacteria create a biofilm that results in significantly less

susceptibility to antibiotics. After the biofilm has matured, *S. aureus* is dispersed to the rest of the body via the circulatory system where it can cause a host of secondary diseases.[30]

Interestingly, variations within *fnbpA* have been found in isolates patients who develop *S. aureus* endocarditis, but not in isolates from similar patients who are colonized with the bacteria without developing invasive disease. These polymorphisms have been shown to affect the affinity of FnbpA for fibronectin.[31]

ClfA and ClfB

Clumping factors A (ClfA and ClfB) are a staphylococcal MSCRAMMs of the Sdr family that are important virulence factors and are found in virtually all *S. aureus* isolates. Their major role is to target fibrinogen, but they are capable of binding many other host proteins. ClfA is one of the most thoroughly studied MSCRAMMs and has been shown to be necessary for staphylococcal agglutination of plasma. ClfA mediates of *S. aureus* adhesion to fibrinogen by targeting the 17 amino acids at the carboxy terminus of the fibrinogen γ-chain.[32] ClfB targets residues 316-328 of the fibrinogen Aα chain.[33] Both ClfA and ClfB mediate platelet aggregation in the presence of fibrinogen and play a role in infective endocarditis. ClfA has been shown in animal studies to play critical roles in infective endocarditis, septic arthritis, kidney abscess formation and

septic bacteremia. In addition to proving efficacious in animal models, antibodies against ClfA have shown some promise as treatment in humans.[34]

Both ClfA and ClfB display additional mechanisms of virulence through the targeting of additional host proteins. ClfA has been shown to target Factor I, a regulator of the alternative complement pathway, resulting in reduced local activation of the host innate immune system.[35] Another important mechanism by which ClfB benefits the bacteria is in the colonization of humans via its binding of Cytokeratin 10.[36,37]

SdrG

SdrG is an MSCRAMM of the Sdr family that is only found in *Staphylococcus epidermidis* that has been discovered to bind to fibrinogen. SdrG binds to the residues 6-20 of the fibrinogen Bβ chain, an especially important site because this is the thrombin cleavage site in the Bβ chain. The binding of SdrG to fibrinogen inhibits cleavage by thrombin, resulting in an inhibition in the release of fibrinopeptide B and the generation of fibrin for clot formation.[38] Like *S. aureus, S. epidermidis* is a bacterium that colonizes humans and is an opportunistic pathogen in humans. It has neither the same set of MSCRAMMs as *S. aureus,* nor does it does it display the same degree of antibiotic resistance that MRSA has shown. For these reasons, *S. epidermidis* does not cause the variety nor severity of pathology of *S. aureus*.

However, *S. epidermidis* does have the ability to cause invasive disease. It is most noted for its ability to cause biofilm formation and act as an opportunistic pathogen in catheters and medically implanted devices.[39] SdrG contributes to this virulence through its high affinity interaction with fibrinogen. Furthermore, the similarity of the ligand binding domain of SdrG to MSCRAMMs in *S. aureus* has helped to elucidate the structural mechanisms by which MSCRAMMs of the Sdr family are able to bind to their target ligands.[38]

SdrCDE Gene Cluster

Three members of the Sdr family of MSCRAMMs occur next to each other in the core genome of *S. aureus*. These three MSCRAMMs were named s*drC*, s*drD* and *sdrE* to reflect their proximity in what was originally assumed to be an operon. A recent publication has shed light on the regulation of these genes and showed them to be independently regulated, strongly suggesting that this is a gene cluster instead of an operon. Specifically, it was shown that *sdrD* transcription occurred in response to different stimuli than *sdrC* or *sdrE.* [40]

The three MSCRAMMs in this gene cluster share a similar domain organization. Starting from the amino terminus, each of these MSCRAMMs with a 30 amino acid signal sequence, followed by the ligand binding A domain. Like ClfA and ClfB, the A domains of these three MSCRAMMs are divided into the N1, N2 and N3 domains. Unlike ClfA and ClfB, each of these MSCRAMMs contains

3-5 B-repeats. Specifically, SdrC contains 2 B-repeats, SdrD contains 5 B-repeats, and SdrE contains 3 B-repeats. The B-repeats function mainly to further protrude the ligand binding domain from the surface of the bacteria, but there is evidence to suggest that these domains are capable of low affinity interactions with the extracellular matrix as well. The B-repeats in SdrF, an Sdr family member from *Staphylococcus epidermidis*, have been reported to be able to type I collagen.[41] Additionally, the SdrD B-repeats were shown to bind calcium. These data also suggest that the binding of calcium brought greater structure to the B-repeats.[42]

Like ClfA and ClfB, there is an R region in each of these three MSCRAMMs that is made up of SD repeats that span the cell wall. Finally, there is a canonical LPXTG that allows sortase to anchor these proteins to the surface of the bacteria. The MSCRAMM Bbp was discovered and named through a separate series of experiments, but we propose that it is actually an allelic variant of SdrE based on numerous factors covered below (Chapter 2). It has the same domain organization as SdrE and is highly similar by amino acid sequence.

Interestingly, the gene frequencies of the genes of this gene cluster are not consistent. *sdrC* is present in almost all isolates, similar to *clfA* and *clfB*. In contrast, *sdrD* and *sdrE* show variable frequencies depending on the type of infection from which the isolates are sampled.

It appears that each of the proteins in the cluster have a different function. The Höök Lab has published that SdrC binds to β-neurexin, a protein found on the surface of cells of the central nervous system that plays a role in cell-cell adhesion and signaling.[43] *S. aureus* is able to cause brain abscesses, although this usually represents a secondary infection. The nearly ubiquitous presence of *sdrC* and the fact that most MSCRAMMs bind to multiple ligands suggest that there are other ligands for SdrC.

While a specific ligand has yet to be found for SdrD, multiple reports suggest that this MSCRAMM plays a role in attachment to the epithelium. Heterologous expression in the non-pathogenic bacterium *Lactococcus lactis* shows that the presence of *sdrD* allows for bacterial attachment to desquamated epithelial cells.[44] Desquamated epithelial cells are the most common cell type in the anterior nares, a location where *S. aureus* has been found to be capable of consistently or intermittently colonizing. *sdrC*, c*lfB* and i*sdA* were also found to play a role in attachment to this tissue type. Further epidemiological evidence suggests that SdrD also plays a role in staphylococcal pneumonia.

Until recently, *sdrE* had no known function. It has been suggested that it plays a role in the indirect activation of platelets. Also, molecular epidemiology studies suggest that it plays a role in invasion of the host. Recently, it was published that SdrE interacts with Factor H, a regulator of the alternative complement pathway that reduces the local activation of the innate immune system. With the established binding of ClfA to Factor I, a protein to which

Factor H binds *in vivo,* it is not surprising that *S. aureus* has multiple virulence factors that target multiple components of this pathway. Specifically, there is a clear benefit to be gained by appropriating the immunomodulatory functions of Factor I and Factor H in order to blind the host's innate immune system to the bacteria.[45]

Bbp was discovered in a search for surface proteins on *S. aureus* that play a role in attachment to bone tissue in osteomyelitis. Further work found that this MSCRAMM bound specifically to the bone extracellular matrix component bone sialoprotein.[46] Additionally, epidemiological studies show that *bbp* is overrepresented in staphylococcal osteomyelitis isolates. Our lab has recently shown that Bbp also binds to human fibrinogen, specifically to residues 561-575 of the fibrinogen Aα chain. [47]

Structure-function relationships in the Sdr family of MSCRAMMs: The dock, lock and latch model

One of the defining characteristics of the Sdr family of MSCRAMMs is the display of a ligand binding domain away from the surface of the bacterium via a stalk like region. This ligand binding domain is made up of two Immunoglobulin-like Folds (Ig-like Folds) that are linked by a Linker domain to form a ligand binding trench.

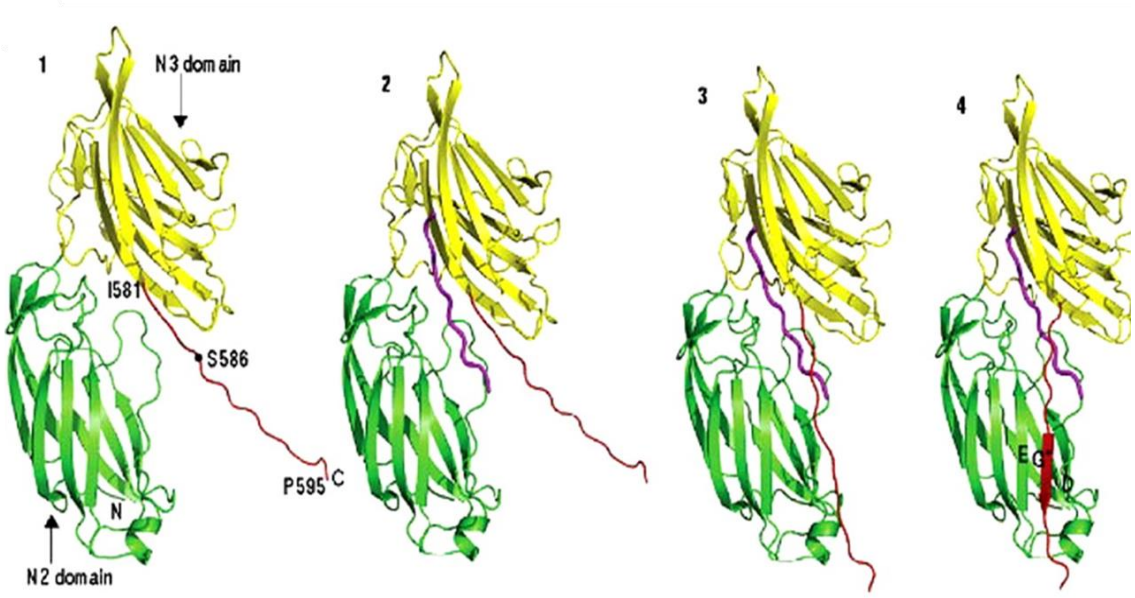


Figure 1-3. Dock, Lock and Latch Mechanism. (G. Vannakambadi)

Like the prototypical Ig-like fold, each of the Ig-like folds in the ligand binding domains of these MSCRAMMs are made up of β-sheets.[48] This secondary structure is confirmed by Circular Dichroism, where the distinctive profile of a β-sheet is seen, and from solved crystal structures of MSCRAMMs like ClfA, ClfB and SdrC. MSCRAMMs of the Sdr family frequently bind to their target ligands via the Dock, Lock and Latch model.[49]

In Figure 1-3 provided by collaborator Dr. Vannakambadi, the Dock, Lock and Latch model is shown with SdrG and a peptide representing the target binding site of SdrG on human fibrinogen. In Panel 1, the N2 and N3 domains of SdrG are shown in the open conformation connected by the linker domain and

the Lock-Latch domain shown in red. These domains form the ligand binding trench, which is where the peptide ligand forms the majority of its contact points with the MSCRAMM (Panel 2). After the ligand is bound in the trench, the flexible Lock domain moves over the ligand and binds to it, placing the MSCRAMM in the closed conformation. Finally, the Latch domain inserts into the N2 domain via β-strand complementation, fixing the MSCRAMM in the closed state. At this point, the MSCRAMM ligand binding domain is wrapped around the ligand while the ligand enters and exits through the trench in the middle of the molecule.[38]

CLASSIFYING AND MONITORING *Staphylococcus Aureus*

Previously, *Staphylococcus aureus* isolates that caused infections were collected and the categorized based on Pulse Field Gel Electrophoresis (PFGE). This approach relies on the migration of enzymatically digested *S. aureus* genomic DNA through an acrylamide gel. Isolates of different lineages show distinct, repeatable banding patterns. PFGE was replaced by Multi-Locus Sequence Typing (MLST), wherein 450-500 bp fragments of important housekeeping genes are sequenced. In *S. aureus,* a minimum of seven housekeeping genes are used to classify the isolate based on the public database. For isolates that are very similar according to PFGE or MLST analysis, *spa* typing and SCC*mec* typing can provide an increased degree of resolution.[50]

The increasing efficiencies of DNA sequencing, both in terms of speed and cost, have resulted in a rich data set of *Staphylococcus aureus* sequence data becoming available over recent years. Where molecular epidemiology previously looked at the presence or absence of an established, relatively small set of genes, there are exciting new possibilities in the field to track changes in nucleotide sequence on a much different scale.

One way this technology has provided new information is the study of whole genome sequences including accessory genomes or mobile genetic elements. For example, in Price *et al,* 85 *S. aureus* CC398 isolates were gathered from a variety of human and animal hosts and then had their whole genomes sequenced. These data were analyzed using single nucleotide polymorphism (SNP) analysis to create an ancestral tree of the isolates. Importantly, they were able to show the one of the most common clonal complexes that causes Livestock Acquired – Methicillin Resistant *S. aureus,* CC398, originated as a human isolate that was susceptible to methicillin. They were able to trace changes that show adaptation to the animal hosts and resistance to a wide range of antibiotics including methicillin.

Another way in which whole genome sequencing has proven useful is to more fully understand the movement of mobile genetic elements such as prophages. The prophage φSa3 has been found to contain virulence factors that play a role in modulating the human innate immune system; it has also been shown to disappear from *S. aureus* isolates that have jumped from humans into

an animal host. The prophage φAvβ3 has been discovered in MSSA CC398 isolates taken from turkeys.[20] This prophage contains genes that are projected to play a role in virulence specific to avian species. Another example of a gain of functionality is the *sasX* gene, which was recently discovered in a study of MRSA infections in China and represents a new putative adhesion molecule that plays a role in virulence. [51]

In addition to being able to observe gains in functionality in the accessory genome, whole genome sequencing also allows to detect for loss of function mutations in virulence factors in the core or core variable genome of *S. aureus*. For example, studies have noted the accumulation of stop codons in human-specific genes, such as *clfA, clfB,* and *hlaB,* in isolates taken from animal species such as cows.[25] While these changes could be monitored through PCR and sequencing of the specific genes, whole genome sequencing allows to sort through all of the genes to look for changes.

The growing, powerful capabilities for whole genome sequencing of pathogens have resulted in observations regarding which virulence factors play putative roles in differing types of infection. In the case of *S. aureus,* especially MRSA, these observations have allowed for greater understanding of antibiotic resistance, host adaptation, and adhesion to implanted medical devices.

CHAPTER II

MATERIALS AND METHODS

COMMERCIAL REAGENTS

Phusion DNA Polymerase, BamHI, HindIII, Dpn1, T4 Ligase, and their requisite buffers were purchased from New England Biolabs. Luria Broth powder, agar, imidazole, SigmaFAST OPD kits were purchased from Sigma Aldrich. Isopropyl-β-D-thiogalactopyranoside (IPTG) and ampicillin was purchased from Gold Bioscientific. 5 mL HisTrap HP, 5 mL HiTrap Q, and Sephacryl S-200 16/60 columns were purchased from GE Healthcare. 0.2 μm Filters and 10K MWCO Concentrators were purchased from Millipore. Medium-binding, flat bottom, 96-well Microlon plates were purchased from Phenix Research. Human and mouse fibrinogen was purchased from Enzyme Research. Cat, cow, dog, pig and sheep fibrinogen were ordered from Sigma. Superblock (in Tris-Buffered Saline), Dialysis Cassettes, were purchased from Pierce-ThermoScientific. Goat-anti-Rabbit-Horseradish-Peroxidase was purchased from BioRad. Bovine serum albumin fraction V (BSA) was from Serological Proteins, Inc. Peptides were ordered from Biomatik and the HanHong Group.

Abs280, Abs260/280, 96-well plates were read using a SpectraMax5 (Molecular Devices). Surface Plasmon Resonance was performed by Xiaowen Liang using a Biacore 3000 (GE Healthcare). Isothermal Titration Calorimetry

was performed using a VP-ITC (Microcal/GE Healthcare). The data were analyzed using Origin 5.0 software. Nucleotide and Amino Acid sequences were analyzed and phylogenetic trees were made using Mega5 software.

METHODS

Sequence analysis, phylogeny and dendograms

Nucleotide sequences of s*drE/sdrE1* and *bbp/sdrE2* were gathered from collaborators and the public database Patric. Sequences were aligned and analyzed using Mega5, which was also used for phylogeny analysis and dendogram construction.

Recombinant N2N3 constructs

The N2N3 regions of Bbp and SdrE were cloned according to previously described protocols. Briefly, *S. aureus* genomic DNA was purified from isolates containing the requisite genes. Primers were made that allowed the amplification of the N2N3 region in PCR. The PCR product was digested with nucleases BamHI and HindIII, then ligated into pQE30 vector that was digested with the same enzymes.

Table 2-1. List of peptides used in these studies.

#	Sequence	Comment
1	SKQFTSSTSYNRGD	Fibrinogen Aα 561-574
2	SKQFVSSTSYNRGD	FgAα-T565V
3	SKQFTSSTSYYRGD	FgAα-N571Y
4	SKQFTSSTSYNGGD	FgAα-R573G
5	LKRFPVQGG	Bone Sialoprotein
6	LKRFTSSTSYNGRGD	BSP-FgChimera
7	SYSKQFTSSTSYNR	Fg Shift (569-572)
8	LKRFPVQGSSDSS	Bone Sialoprotein Long
9	LKRFTVQGSSDSS	BSP-Fg Point mutation
10	LKRFTVQGSSDSR	BSP-Fg Point mutation
11	LKQFTSSTSYNGR	BSP-Fg Point mutation
12	SKRFTSSTSYNGR	BSP-Fg Point mutation
13	SKQFVTSSTTYNRGD	Dog Fg
14	SKQFVSSSTTVNRGG	Cow Fg
15	SKQLVATSKTYNRGDS	Cat Fg
16	SKQTITKTINREGR	Pig Fg
17	SHNMTTTLNYRDG	Factor H Peptide 1 - SCR 13

The ligated vector was then transformed into *E. coli* strain XL1-Blue (Stratagene) for long term storage. The Bbp and SdrE constructs were made by Jenny K. Horndahl and Vanessa Vazquez, respectively, in our laboratory.

Prior to each expression, the vector was transformed into the *E. coli* TOP 3 (Invitrogen) cell line. Subsequently, overnight starter cultures were made from a colony taken from the transformation plate, which were then used to inoculate 1 liter samples of Luria Broth Media containing 100 μg/mL Ampicillin. After these cultures reached exponential phase, they were cooled for 30 minutes at 4°C. Next, cultures were induced with 500 μM IPTG, then incubated at 18°C for 24-72 hours with shaking. Cultures were then pelleted, washed with 1x PBS, and stored at -80°C.

After thawing, cell pellets were re-suspended in 1x TBS then lysed in a French Press or Cell Disruptor, pelleted in an ultracentrifuge, filtered and then applied to a Ni^{2+} HisTrap Affinity column. Gradient elution was accomplished with imidazole. Peak fractions were collected, dialyzed into 25 mM Tris (pH 7.9) buffer with no salt, then applied to a HiTrap Q Column, where the protein of interest was eluted with an increasing gradient of NaCl. To assess the purity of resulting protein sample, peak fractions were run on 10% SDS-PAGE and stained with Coomassie.

SdrE2/Bbp-N2N3 point mutant and chimeric constructs

Site-Directed Mutagenesis was performed on the SdrE2-N2N3 constructs to create a number of point mutants. Briefly, purified *sdrE2-N2N3* in pQE30 was used in a PCR reaction that included forward and reverse primers that contained the desired mutation. These primers bound to the wild type construct at the desired site of mutation. PCR amplification of the mutated plasmid and a subsequent treatment of the PCR product with DpnI removed the starting DNA material. The resulting Dpn1-treated PCR product was digested with BamH1 and HindIII and then ligated into pQE30. The construct was confirmed by sequencing and transformed into *E. coli* XL1-Blue for long term storage.

A modified form of Overlap Extension PCR was used to create SdrE1-LockChimera and SdrE2-LockChimera. Since the Lock region is near the end of the N2N3 constructs, a series of primers were designed that would allow for replacing the Lock domain with the corresponding Lock domain from the other allelic variant. The first primer pairs with the 15 nucleotides before the Lock domain and contains an additional 15 nucleotides with the new desired Lock region that will not bind to the parent construct. PCR performed with this primer and the wild type forward primer resulted in a PCR product that contained a truncated N2N3 construct with 15 new nucleotides at the end. The next primer paired with the newly mutated nucleotides and also contained a 15 nucleotide overhang corresponding to the desired sequence. A third primer was needed to complete the construct. The same forward primer was used in each of

the 3 PCR reactions. The resulting PCR product was digested with BamH1 and HindIII and then ligated into pQE30. The construct was confirmed by sequencing and transformed into *E. coli* XL1-Blue for long term storage.

The same purification protocol was used for mutant constructs.

Rabbit-anti-BbpN2N3 antibodies

Polyclonal antisera to BbpN2N3 were generated in two rabbits at Rockland Immunochemicals under the Fast Production Protocol. IgG was purified using protein A-sepharose (ThermoFisher) affinity chromatography. Next, the IgG was cleared for crossreactive binding to Sdr proteins before positive affinity purification on BbpN2N3 coupled to EZlink beads (ThermoFisher).

Animal fibrinogen ELISA-type assay

Binding of SdrE1, SdrE2 and mutant/chimeric constructs of these MSCRAMMs to fibrinogen from various animals was first determined using an ELISA-type solid-phase assay. Fibrinogen from humans, dogs, cats, cows, pigs, sheep, and mice along with BSA were coated on microtiter wells at 1 μg per well overnight at 4°C in bicarbonate (0.1 M $NaHCO_3$) buffer (pH 8.0). The next day, the wells were washed with TBS-T (0.1% Tween-20) and blocked with Superblock (TBS) for 3 hours at room temperature. Wells were incubated with

dilutions of the MSCRAMM (0.1-10 μM). Detection of binding was achieved by Rabbit anti-SdrE1-N2N3 or Rabbit anti-SdrE2-N2N3, followed by Goat anti-rabbit HRP. Finally, the wells were developed with SigmaFast OPD and the absorbance was measured at 450 nm with a SpectraMax M5 plate reader (Molecular Devices).

Peptide inhibition ELISA-type assay

Peptide Inhibition ELISA-type Assays are performed similarly to the Animal Fibrinogen ELISA-type Assay. Wells are coated overnight with 1 μg per well at 4°C in bicarbonate (0.1 M $NaHCO_3$) buffer (pH 8.0). The next day, the wells were washed with TBS-T (0.1% Tween-20) and blocked with Superblock (TBS) for 3 hours at room temperature. During blocking, 0.5 μM of the designated MSCRAMM was incubated with increasing concentrations of the designated peptide (2.5 – 50 μM) for 1 hour. Peptides represent the human fibrinogen target sequence and the corresponding residues from fibrinogen from species including dogs, cats, cows, and pigs.

Subsequently, the MSCRAMM-peptide mixture was added to the coated and blocked microtiter wells. Detection of binding was achieved by Rabbit anti-SdrE1-N2N3 or Rabbit anti-SdrE2-N2N3, followed by Goat anti-rabbit HRP. Finally, the wells were developed with SigmaFast OPD and the absorbance was measured at 450 nm with a Thermo Max plate reader (Molecular Devices).

Decreased signal is interpreted as the MSCRAMM binding to more peptide and thus being unable to bind to the coated wells.

Isothermal titration calorimetry

Isothermal Titration Calorimetry (ITC) was performed with a VP-ITC (MicroCal) with SdrE1-N2N3 or SdrE2-N2N3 and previously mentioned peptides. The MSCRAMMs were dialyzed into TBS (1x) over 48 hours at 4°C and their UV absorbance was measured at 280 nm to in order to accurately and directly measure their concentrations. Final dilution of the MSCRAMM to 15 μM was performed with buffer from the last round of dialysis. Peptides were diluted from a stock concentration of 10 g/L (~6 μM) to their final concentration (100-250 μM) using the same buffer from the last round of dialysis. Peptide was placed into the syringe and injected into the cell containing the MSCRAMM in 10 μL intervals every 300 seconds after an initial 5 μL injection. Experiments were performed at 30°C with stirring set at 300 rpm, reference power set at 15 and feedback gain on high. Data was analyzed using Origin 5.0 (MicroCal) software.

Surface plasmon resonance

Surface Plasmon Resonance (SPR) was performed by Xiaowen Liang using a Biacore 3000. Two sensor chips were coated with fibrinogen from multiple species. The 4 channels of the first chip were coated with human, cow and pig FG as well as a control. The 4 channels of the second chip were coated with dog, cat, and sheep fibrinogen and the control. SdrE1-N2N3 and SdrE2-N2N3, dialyzed into TBS, were injected over the chips for 120 seconds, followed by buffer. The K_{on} and K_{off} were measured. BIAevaluation software was used to analyze the data and generate graphs.

CHAPTER III

EPIDEMIOLOGY OF ALLELIC VARIANTS *sdrE/sdrE1* AND *bbp/sdrE2*

INTRODUCTION

Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMMs) are an important class of virulence factors that are found in many Gram-positive organisms, including the human pathogen *Staphylococcus aureus*. These surface proteins have a variety of functions including host adhesion and immune evasion. Among the staphylococcal MSCRAMMs is a family of proteins termed the Sdr Family. These molecules derive their name from the Serine-Aspartate (SD) Repeats that all members contain. As shown in Figure 3-1, the members of this family have similar subdomain architecture and include important MSCRAMMs such as *clfA*, c*lfB* and the *sdrCDE* gene cluster. Intriguingly, the *bbp* gene shows high sequence similarity to the Sdr family members, specifically *sdrE*.

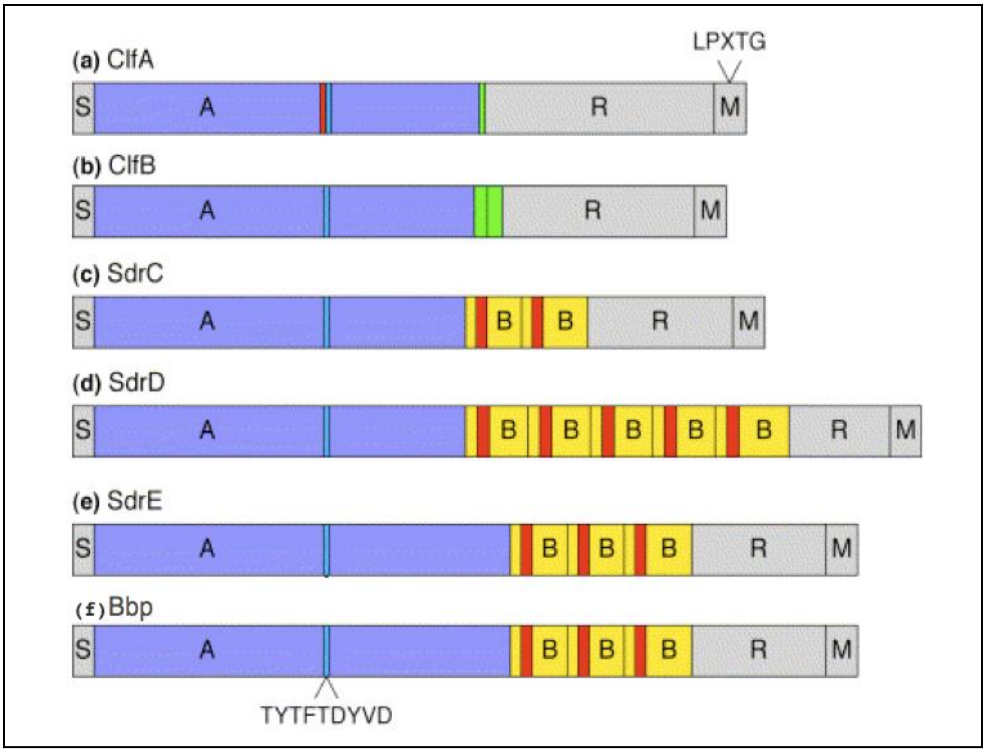


Figure 3-1. Schematic of Sdr Family of MSCRAMMs. (V.Vazquez)

In 1989, it was published by Rydén *et al* that *S. aureus* isolates from osteomyelitis patients are able to bind bone sialoprotein (BSP).[52] Further work by these investigators isolated an MSCRAMM as the bacterial surface factor responsible for this binding. They termed this MSCRAMM Bone sialoprotein–Binding Protein, Bbp.[53] The interaction between Bbp and BSP was further isolated to the nonapeptide sequence, LKRFPVQGG, that occurs in the N-terminal half of BSP.[46] Additional evidence for the proposed role of Bbp in staphylococcal osteomyelitis is seen in molecular epidemiological studies that

show Bbp to be overrepresented in osteomyelitis strains compared to staphylococcal strains from all pathologies. [54]

Our lab recently published that Bbp also binds to human fibrinogen.[47] While many MSCRAMMs are known to bind more than one host extracellular matrix (ECM) protein, biochemical screens show that fibrinogen is the only ECM protein targeted by Bbp. These screens are ELISA-type assays in which microtiter wells coated with ECM components are probed with increasing amounts of recombinant Bbp.

Until recently, the function of SdrE remained unknown. Epidemiological studies suggested that *sdrE* might play a role in invasiveness based on an increased frequency in invasive isolates as compared to carriage isolates, but offered no predictions regarding ligand or mechanism. Recently, it was published that SdrE binds to Factor H, a key regulator of the Alternative Pathway of the human complement system.[45] The ability of *S. aureus* to coat itself with a complement inhibitor represents a potent putative virulence mechanism. This capability has been demonstrated in other bacterial and yeast pathogens through virulence factors that specifically interact with Factor H.

Despite the disparate origins of discovery for SdrE and Bbp, we propose here that these are not separate MSCRAMMs but instead are allelic variants. The striking similarities between the amino acid sequences have led some laboratories to consider Bbp and SdrE the same virulence factor, while others

consider them allelic variants. However, the vast majority of the field still considers Bbp and SdrE to be independent virulence factors. Interestingly, when Bbp and SdrE are treated as allelic variants, their gene frequencies can be analyzed for clues as to any potential roles in virulence. The gene frequencies of these allelic variants in both animal and human staphylococcal isolates are compared and sequence variations are studied.

RESULTS

Bbp and SdrE are allelic variants

With the technical capabilities of genome sequencing expanding at a rapid rate, there is much richer sequence data available for *S. aureus,* including whole genome sequences. Analysis of the available data shows many pieces of evidence that support the hypothesis that *sdrE* and *bbp* are indeed allelic variants rather than separately expressed virulence factors. First, *sdrE* and *bbp* genes are found together in less than 1% of *S. aureus* isolates. Furthermore, these two genes occur in the same location of the core genome of *S. aureus*. This location occurs after *sdrC* and *sdrD* in the *sdrCDE* gene cluster, when both *sdrC* and *sdrD* are present.

Beyond chromosomal organization, evidence for the allelic variant hypothesis exists at the protein domain architecture of *sdrE* and *bbp*. Both proteins are members of the Sdr subfamily of MSCRAMMs. All members of this

subfamily contain similar subdomain architecture with a ligand binding A domain presented from the surface of the cell by a stalk-like repeat (R) region and a variable number of B Repeats. Both *sdrE* and *bbp* contain A domains followed by 3 B-repeats, with the size of the aligned subdomains being equivalent. The SD-repeat containing R region is variable in length between *sdrE* genes and between allelic variants. The size of the genes correlates directly to the number of SD-repeats and is not indicative of variations within the gene or within allelic variants.

Phylogenetic analysis reveals that *bbp* and *sdrE* are more similar to each other than the other members of the Sdr subfamily (Figure 3-2). Moreover, they are more similar to the neighboring *sdrC* and *sdrD* of the *sdrCDE* gene cluster than they are to *clfA* or *clfB*. This suggests, that the *sdrCDE* gene cluster shares a common ancestor and that the individual genes likely arrived via gene duplication events. The locations of these genes suggest that first with *sdrD* formed from *sdrC,* followed by *sdrE* formation.

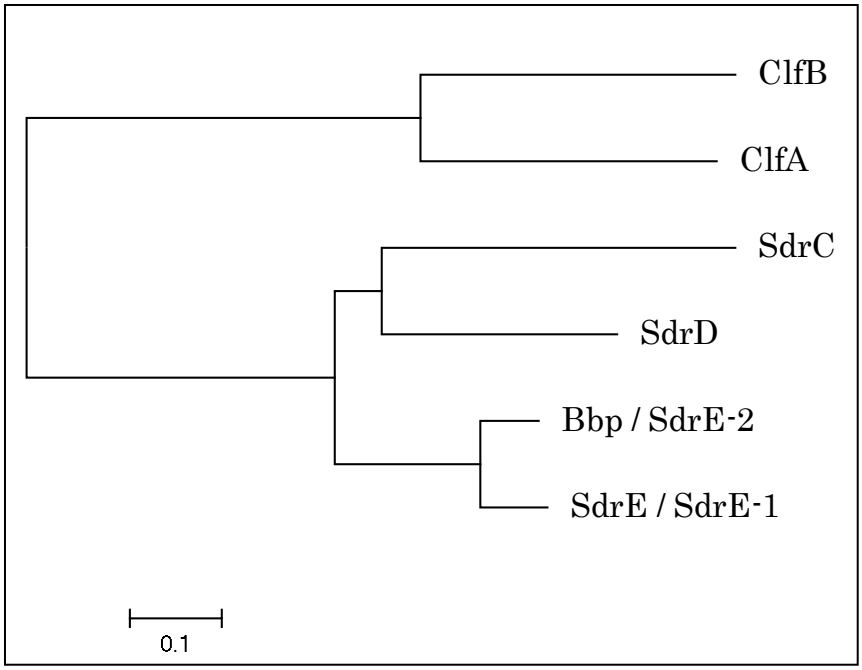


Figure 3-2. Dendogram showing similarity in Sdr family of MSCRAMMs

sdrE and *bbp* are highly similar; they are 87% identical by amino acid sequence. Interestingly, there are distinct differences in the rate of variation depending on the subdomain. Within the ligand-binding N2N3 domains, *sdrE* and *bbp* display a more modest 67% sequence identity. However, in the N1 subdomain and B repeats, there is a 95% sequence identity (Figure 3-3).

If *sdrE* and *bbp* were unrelated virulence factors that simply showed high sequence similarity, the variations would be expected to occur at similar rate across the entire gene. Instead, variations are relatively concentrated to 2

subdomains, N2 and N3, suggests that these two genes are allelic variants that originated from a similar ancestor. Mutations likely accumulated in one variant that conferred a competitive advantage in at least some of the different host environments that *S. aureus* encounters.

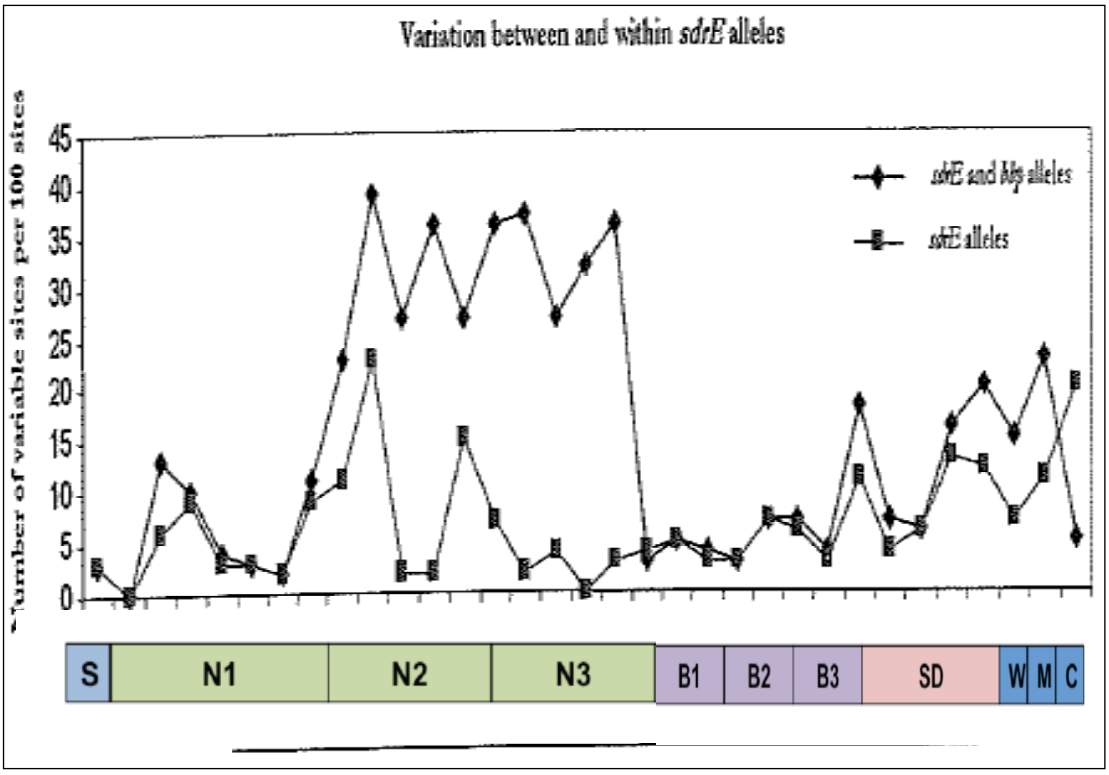


Figure 3-3. Locating regions of high variation in *sdrE*.(Adapted from Ed Feil, "Single Nucleotide Polymorphisms in Bbp", University of Bath, 2003)

It is surprising to see this distribution of variations between *sdrE* and *bbp* given that the N2N3 domains are responsible for the effector function of the MSCRAMMs. It is more common for mutations to occur in regions whose putative roles involve support or architectural purposes while conserving residues that are integral for binding or enzymatic function. One possible explanation for this observation is that one variant accumulated enough mutations that the end result was a gain of functionality in a stepwise manner. Another possibility is that these allelic variants provide a redundant function to *S. aureus*, such that there is less evolutionary pressure to conserve this ligand binding sequence. This possibility appears to be the most likely in light of the published ability of Bbp to bind to fibrinogen with high affinity. *S. aureus* has multiple MSCRAMMs that bind to different regions on fibrinogen with high affinity, including ClfA, ClfB, FnBPA and FnBPB.

While neither *bbp* nor *sdrE* are present at the same level as the ubiquitous *sdrC* individually, it is interesting to note that the percentage of isolates that have either *bbp* or *sdrE* is closer to the rate seen for *sdrC*. Figure 3-4 shows the percentage of isolates that were positive for *bbp* or *sdrE* in panels 1 and 2, respectively. Panel 3 is a combination of data from reports that treated these allelic variants as the same gene and composite totals of the data in panels 1 and 2. Weighted by the number of isolates used in each study, *bbp* is found in 31.8% and *sdrE* is found in 55.5% of the 649 total isolates gathered from various

settings. These percentages in sum, 87.3%, are similar to the reported values from the studies that treated these genes as alleles.[55-59]

Another piece of evidence that *sdrE* and *bbp* are allelic variants is found in a study by Pietro, *et al*¸ of the 4 *agr* loci that sought to find MSCRAMMs and other virulence factors that co-segregated with one or more of the *agr* loci. Their findings included the observations that *sdrE* co-segregated with *agr I* and *agr II*, while *bbp* co-segregates with *agr III*.[57]

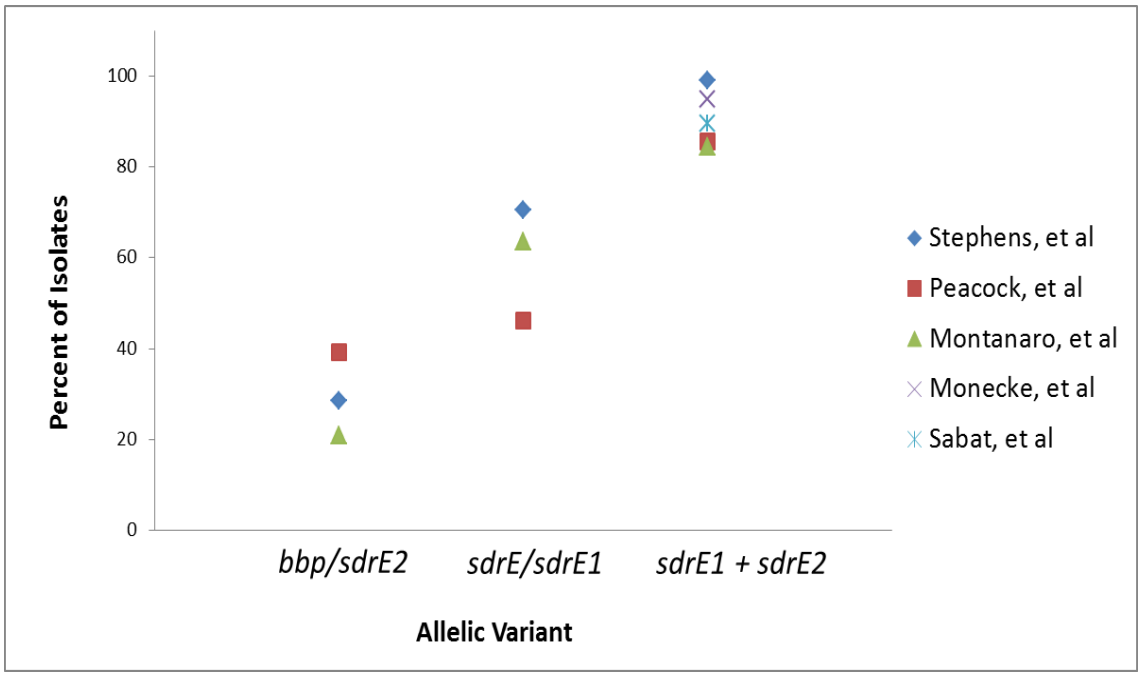


Figure 3-4. Frequency of *sdrE* alleles.

All of these data support the hypothesis that *sdrE* and *bbp* are allelic variants. Furthermore, the current literature contains at least three publications that have cited the high degree of sequence identity as sufficient cause to treat these two genes as allelic variants.[55-57] However, most epidemiological studies still regard them as independent virulence factors.

One of the reasons this is problematic is that the high degree of nucleotide sequence similarity between *sdrE* and *bbp* can result in false positives due to primer overlap. If *sdrE* and *bbp* were widely recognized as allelic variants, *sdrE*$^+$/*bbp*$^+$ isolates would warrant further investigation to check for false positives, gene duplications, mobile genetic elements or other explanations for this rare finding. With the misconception persisting of *sdrE* and *bbp* as individual virulence factors, this scenario for potential error will continue.

In light of these observations, we propose a change in nomenclature. Due to the fact that the location of *sdrE* and *bbp* occurs at the end of the *sdrCDE* locus and that *sdrE* occurs at a higher frequency than *bbp*, we propose changing the name of *sdrE* to *sdrE1* and *bbp* to *sdrE2*. Additionally, *sdrE2* clarifies this gene as a member of the Sdr subfamily. While Bbp was originally named based on its ability to bind Bone sialoprotein, the recent finding that this MSCRAMM binds to fibrinogen raises the need for a name that is not tied solely to one function.

Furthermore, *sdrE* has already been noted in the literature as being a highly polymorphic gene, making the concept of a variant of this gene less surprising.[60] This change in nomenclature would help future studies by reducing confusion between these allelic variants and allow for more accurate epidemiological studies. It would be especially interesting to see if there is a statistically significant difference in the distribution of these allelic variants among different pathologies, locations, and host.

Later we present evidence that there are significant variations in allelic distribution, but there is a large, established publication history of *S. aureus* molecular epidemiology whose data we are not able to fully analyze in depth. With the advances in genome sequencing technology, the expectation is for the available sequence data to rapidly grow in exponential fashion. While it may not be possible to analyze previous data, it is important to change the nomenclature as soon as possible before the high throughput technology is fully brought to bear.

sdrE1 **and** ***sdrE2*** **are allelic variants that form distinct groups upon sequence variation analysis**

The Patric database is a publicly available database of whole genome sequences of *Staphylococcus aureus* isolates. The isolates are gathered from a variety of types of *S. aureus*–induced human pathology and locations across the

world. We searched this database to identify the isolates that contained *sdrE1* and *sdrE2* and gathered the nucleotide sequences of these genes. Within the isolates, *sdrE1* occurs in 57 of isolates and *sdrE2* occurs in 16 isolates. When the amino acid sequences of *sdrE1* and *sdrE2* were gathered, a majority of the variation was found to occur within the *sdrE1* variant pool.

When these variations are aligned and plotted via Maximum Likelihood Phylogeny Tree, there was no overlap between variants; *sdrE1* and *sdrE2* were completely segregated (Fig 3-5). This method is capable of putative evolutionary ancestors; however, no ancestor could be identified from the publicly available sequences.

In spite of this, there are clues to the gene ancestry. A greater amount of variability exists within *sdrE1* relative to *sdrE2*. This suggests that *sdrE1* is the ancestral gene and that *sdrE2* represents an allelic variant that accumulated variations that conferred a strong enough evolutionary advantage that any intermediate variants were outcompeted by either *sdrE1* or *sdrE2*.

The RF122 isolate comes from a bovine *S. aureus* strain and is the only variant sequence seen within *sdrE2* alleles. When the lower amount of variation seen within s*drE2* is viewed in light of the highly specific, high affinity binding of SdrE2-N2N3 with human fibrinogen, it raises the possibility that SdrE2 represents an allelic variant that is highly adapted to the human host.

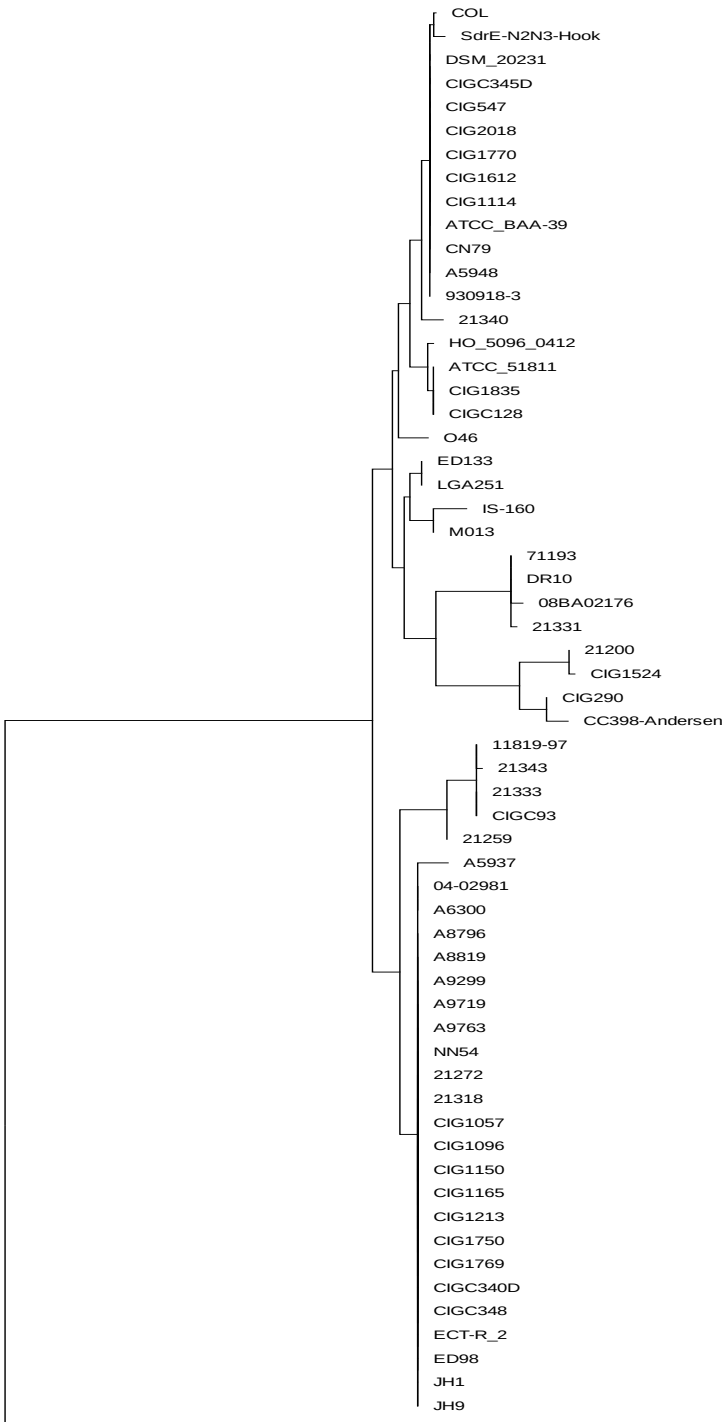


Figure 3-5. Dendogram of *sdrE1* and *sdrE2* from *S. aureus* isolates from humans

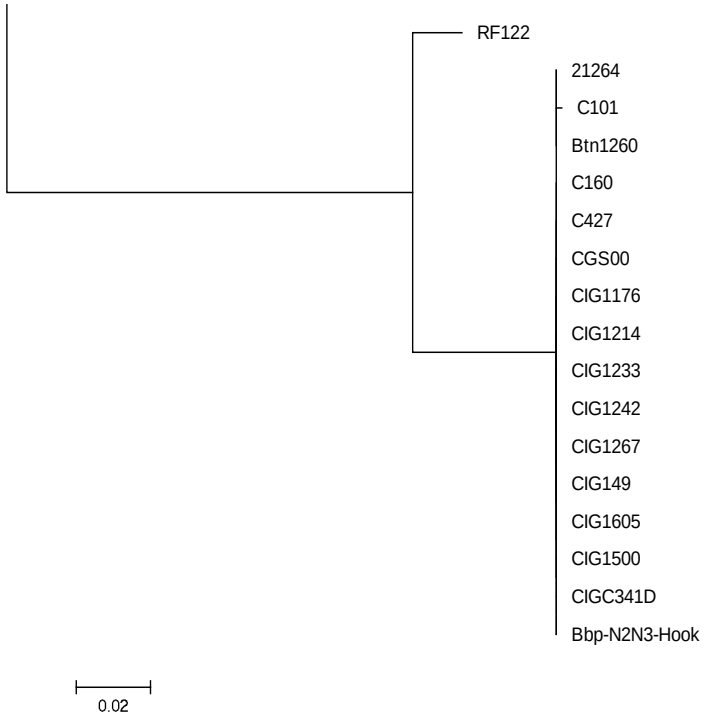


Figure 3-5. Continued

Differences in allelic frequencies in animal and human staphylococcal isolates

To investigate this possibility we expanded our long-standing collaboration with Dr. Ross Fitzgerald. The Fitzgerald lab has acquired a collection of *S. aureus* isolates that were taken from a variety of non-human species. Initial PCR studies suggested that *sdrE1* was overrepresented, while *sdrE2* was dramatically underrepresented when compared to their gene frequencies in human staphylococcal isolates (Fig 3-6). 104 staphylococcal

isolates were gathered from a wide array of species and countries. *sdrE1* was found in 88% of the isolates while *sdrE2* was found in 11% of the isolates. This suggests that the mechanism of virulence of these allelic variants could play a role in species tropism.

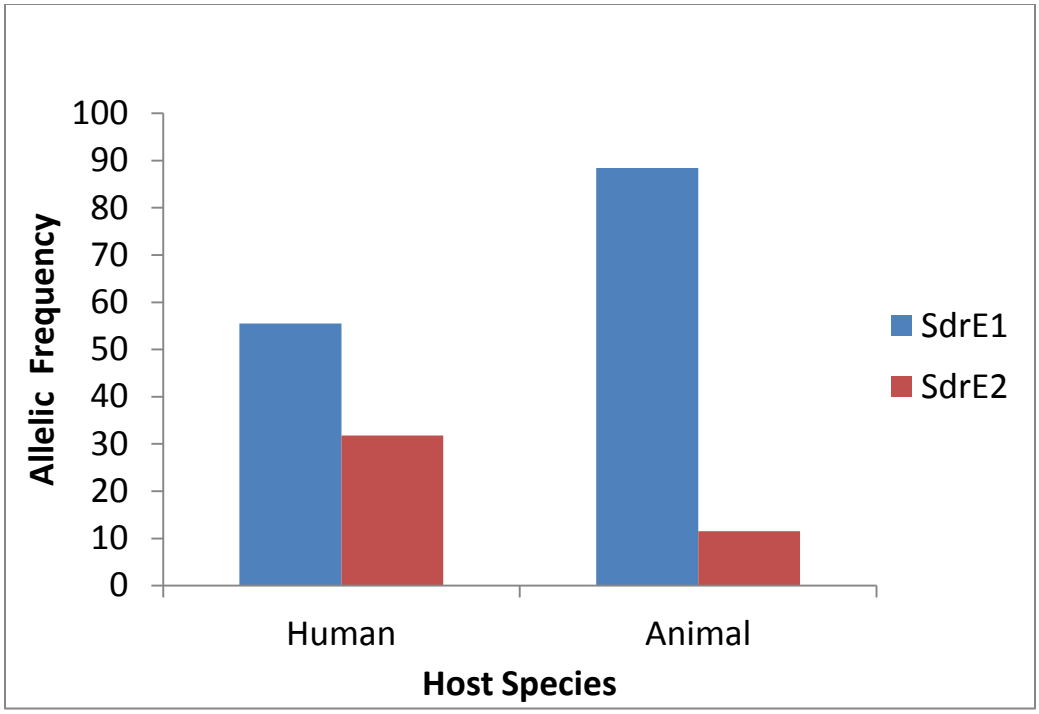


Figure 3-6. Allelic Frequencies in human and animal staphylococcal isolates.

Sequence variations in *sdrE1* and *sdrE2*

The *sdrE1* and *sdrE2* genes were further analyzed through sequencing of the A domains. Figure 3-7 shows a Maximum Likelihood Phylogeny Tree where *sdrE1* and *sdrE2* again segregate completely. This revealed 3-4 major subgroups of *sdrE1* and 2 major subgroups of *sdrE2*. One *sdrE2* subgroup aligns with the RF122 isolate, a known bovine isolate, and consists mainly of other isolates taken from a bovine source. Moreover, 75% of the overall isolates that contain *sdrE2* come from a bovine source. *sdrE1*-containing isolates show a wide species tropism.

DISCUSSION

It is clear from these data that *sdrE* and *bbp* are allelic variants. Based on data from whole genome sequences and sequencing of *sdrE* and *bbp* genes from staphylococcal isolates gathered from around the world, we propose a change in nomenclature to reflect this finding and to prevent confusion. By treating these MSCRAMMs as independent virulence factors, it becomes more difficult to gather data regarding the distribution of these genes and the potential role for host adaptation. We propose that *sdrE* be renamed *sdrE1* and *bbp* be renamed *sdrE2*.

Figure 3-7. Dendogram of *sdrE1* and *sdrE2* from *S. aureus* isolates from animals

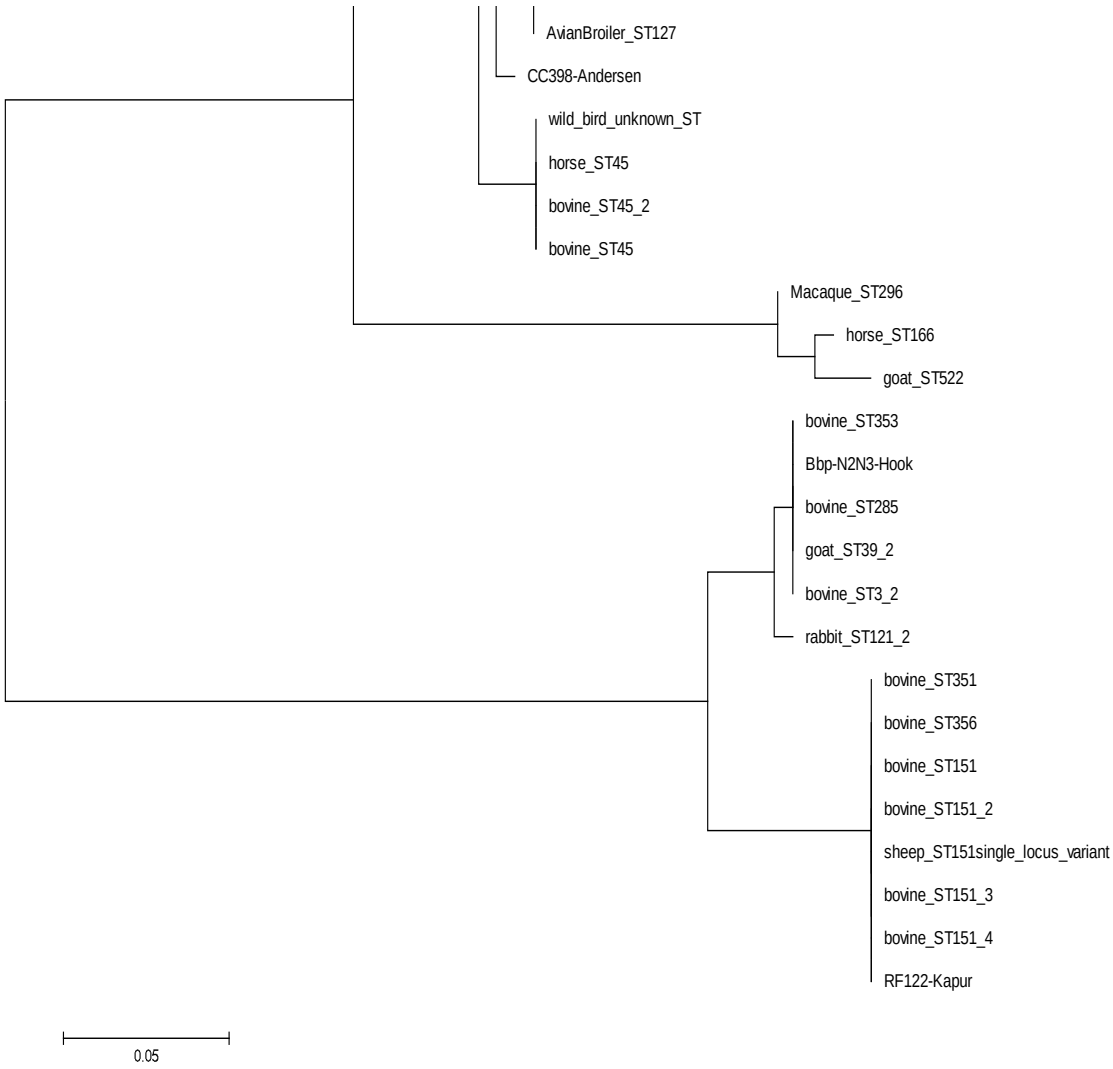


Figure 3-7. Continued

Interestingly, there is a significant difference in the distribution of *sdrE1* and *sdrE2* when animal and human isolates are compared. We hypothesize that the difference in gene frequencies is a reflection of evolutionary adaptation to new animal hosts. *S. aureus* is a human pathogen that is able to survive in animals hosts, but this broad species tropism results in the loss of human-specific virulence factors when the bacteria makes the jump to an animal host. The overrepresentation of *sdrE1* in animal isolates suggests that *sdrE1* provides a greater capacity for virulence than *sdrE2* in animals. The exception to this model is bovine species, where it appears that the two variants have a similar capacity for virulence.

As observed in the staphylococcal isolates gathered from humans, there is a greater degree of variation within *sdrE1* alleles than within *sdrE2* alleles. It is possible that some of these *sdrE1* sub-variants have accumulated mutations that provide a greater virulence in its new animal host than the *sdrE1* sub-variants that occur in staphylococcal isolates from humans.

Of particular interest are variations in *sdrE1* from CC398, a clonal complex that has been identified as one of the primary causes of Livestock Acquired-Methicillin Resistant *Staphylococcus aureus* (LA-MRSA). This clonal complex has been well studied with both whole genome sequencing and other microbiology techniques.

The spread of pathology from this clonal complex in animals makes it likely that it is better adapted to the animal host than other isolates. *SdrE1* represents an opportunity to understand host adaptation at the level of a single gene and potentially at a molecular level.

CHAPTER IV

FIBRINOGEN BINDING PROFILE OF SDRE1 AND SDRE2 GIVES A RATIONALE FOR OBSERVED EPIDEMIOLOGY

INTRODUCTION

Bbp/SdrE2, BSP and fibrinogen

Recently, the Höök lab published that Bbp/SdrE2 binds human fibrinogen. While the MSCRAMM was named for its ability to bind to bone tissue, specifically to Bone Sialoprotein, Bbp/SdrE2 would be one of many staphylococcal MSCRAMMs that have shown the ability to bind multiple ligands. In an extracellular matrix ligand screen done through ELISA-type assay, the ligand binding N2N3 domains of Bbp showed the ability to bind specifically to fibrinogen (Figure 4-1a). In light of the multiple host ligands that are targeted by this MSCRAMM and the previous data indicating that Bbp is an allelic variant of SdrE, we have called for a change in nomenclature where SdrE will now be called SdrE1 and Bbp will now be called SdrE2 . Further experimentation using Far Western Blotting and truncation mutants showed that the high affinity interaction could be to residues 561-575 of the Fg Aα chain (Figure 4-1b).[47]

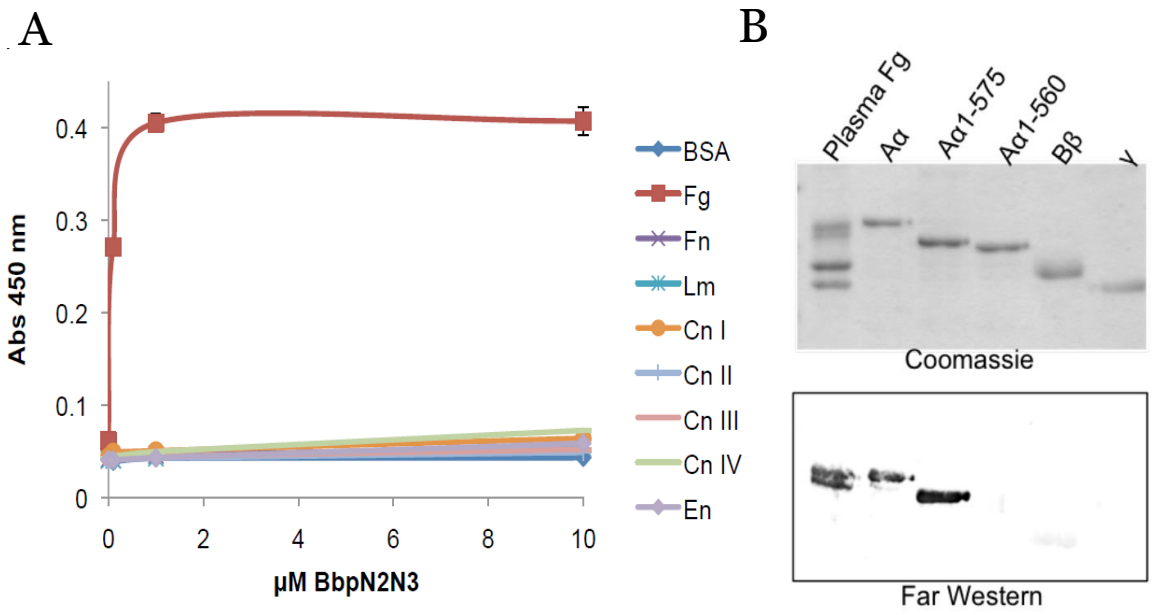


Figure 4-1. Bbp/SdrE2 binds Fibrinogen Aα (V. Vazquez)
(A) Binding of purified N2N3 region of Bbp/SdrE2 to various extracellular matrix proteins. (B) Isolation of the target site of Bbp in human fibrinogen using Coomassie staining and Far Western analysis.

Fibrinogen in hemostasis

Fibrinogen is a major serum protein and acute phase protein produced by the liver that plays a critical role in coagulation as well as roles in inflammation and immune defense. The essential step in coagulation results in fibrinogen being cleaved by activated thrombin to form an insoluble product termed fibrin. In order for prothrombin to be activated to thrombin, a series of enzymatic reactions must first occur (Figure 4-2). This cascade results in an amplification of the initial signal to rapidly ramp up coagulation at the injured site.

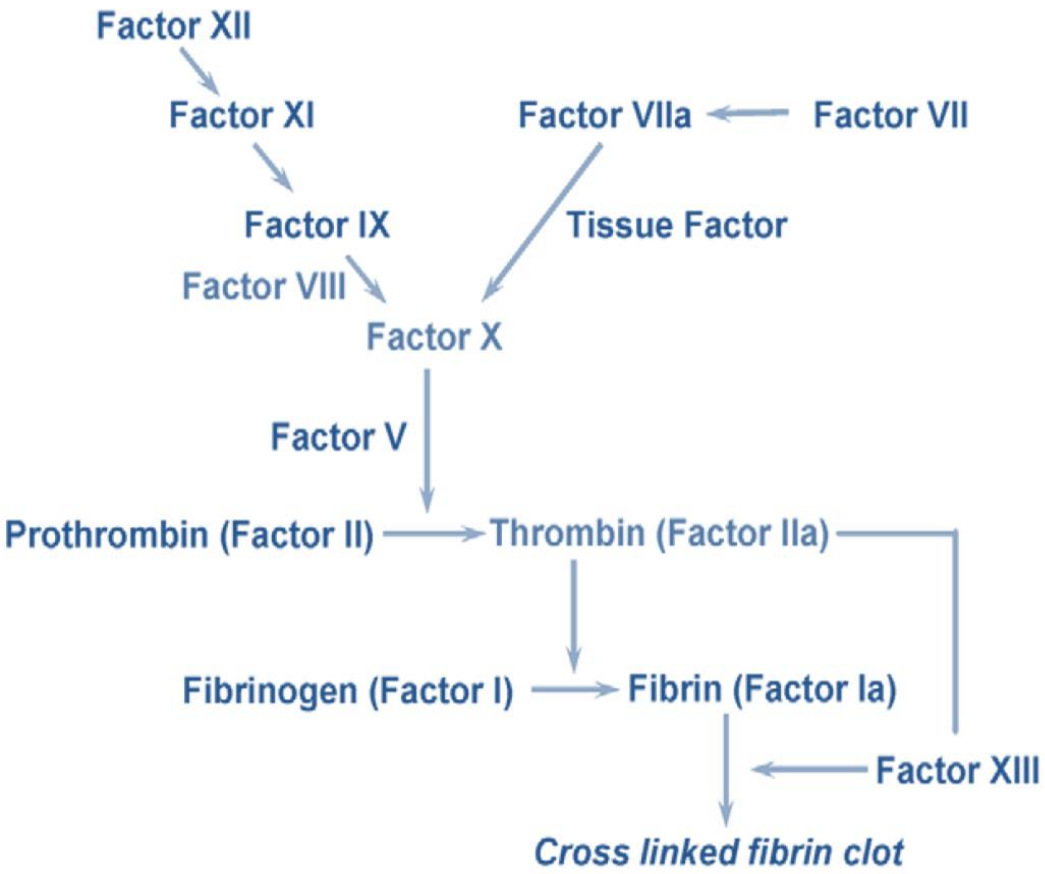


Figure 4-2. Coagulation Cascade

This insoluble fibrin is then cross-linked by Factor XIII to form a mesh-like network that is capable of stopping the flow of blood, attracting pro-coagulation platelets, and recruiting immune cells such as neutrophils. Platelets and neutrophils display integrin molecules on their surface that recognize multiple RGD motifs found within fibrinogen.

Fibrinogen deficiency

While normal blood levels of fibrinogen range from 1.5-3 g/L, fibrinogen deficiencies can occur congenitally or be acquired. Congenital afibrinogenemia is a rare bleeding disorder that occurs at an approximate rate of 1:1,000,000 and encompasses mutations that occur in any of the three genes encoding the three polypeptide chains of fibrinogen.[62] Phenotypically, it presents with spontaneous bleeding or excessive bleeding after minor trauma. The low prevalence of congenital afibrinogenemia is likely a consequence of the importance of fibrinogen in for normal hemostatic function. Acquired fibrinogen deficiency generally occurs as a result of blood loss from trauma, disseminated intravascular coagulation (DIC), cirrhosis or sepsis.

Fibrinogen structure

Fibrinogen is a hexameric molecule that is composed of two sets of three chains (Fig 4-3) that are linked by disulfide bonds. The two sets of chains are connected by a central nodule where the N-termini interact. This region is cleaved by thrombin. Progressing towards the C-termini, the subsequent regions form coiled-coil structures, followed by a globular domain termed the D Domain.[63] After the injured site has healed, fibrinogen is enzymatically cleaved by the activated form of plasminogen called plasmin. During fibrinolysis by plasmin, the D-domains are cleaved. D-dimers are measured clinically as a

diagnostic test for DIC and pulmonary embolism, in which fibrin is being cleaved rapidly.

The αC-domain has been of highlighted in the field of fibrin mechanics where recent findings implicate αC-domain in playing a prominent role elasticity of fibrinogen. Nuclear Magnetic Resonance studies of αC-domain have confirmed that this domain is largely unstructured. In these studies, a hairpin turn was discovered from amino acids 423-450.[64] Moving towards the carboxy terminal of this beta hairpin, within a flexible region of the protein, is the binding site for SdrE.

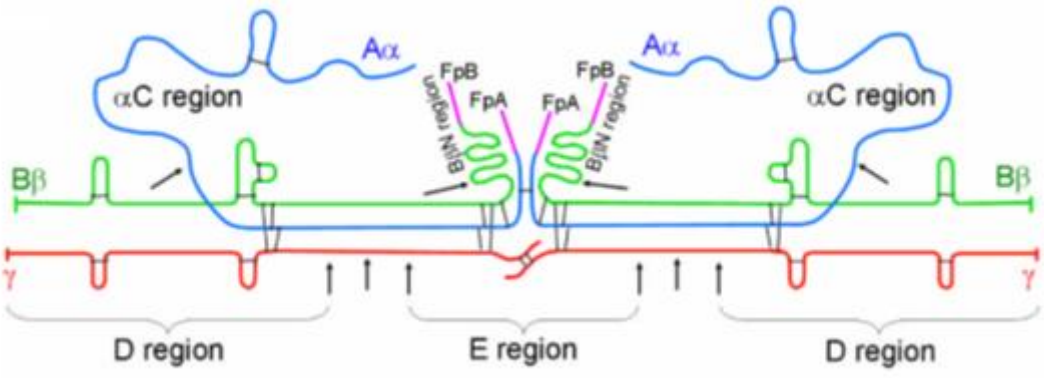


Figure 4-3. Fibrinogen Structure
(A) Individual Fg chains, Aα, (blue) Bβ (green) and γ (red), FpA and FpB: fibrinopeptides A and B; black bars: disulfide bonds; triple arrows: plasmin cleavage sites for D and E fragments; single arrows: cleavage sites for removal of αC and BβN regions. (V.Vazquez)

Biochemical investigation of Bbp-fibrinogen binding

The high affinity binding of SdrE2-N2N3 to the Fibrinogen Aα chain was investigated through an array of biochemical approaches. In Surface Plasmon Resonance experiments, SdrE2-N2N3 showed rapid on and off rates when injected over a human fibrinogen coated chip. A linear dilution series was used to determine a sub-micromolar dissociation constant (K_D = 0.54 μM ± 0.07).

A putative Fg target site was elucidated via ELISA-type solid-phase assays using truncation mutants; from these, a 15 amino acid sequence emerged as the target site of SdrE2-N2N3 on Fg. Subsequently, a peptide was synthesized (Biomatik) that corresponded to this site. This peptide was able to inhibit the binding of SdrE2-N2N3 to Fg-coated wells in ELISA-type solid-phase assay in a dose-dependent manner, showing that this site is the primary target of SdrE2-N2N3 within fibrinogen. Isothermal titration calorimetry was performed with this peptide titrated into a cell containing SdrE2-N2N3. By measuring the heat given off during the exothermic reaction, a K_D = 0.31 μM was calculated.

It was also shown that SdrE2-N2N3 interacts minimally with fibrinogen from other animals. This is especially noteworthy in light of the high degree of homology in this region of fibrinogen between different species.

Putative role in pathogenesis for Bbp-fibrinogen interaction

The widespread presence and multiple physiological roles of fibrinogen in humans make it an excellent target for the opportunistic pathogen *Staphylococcus aureus*. One example of importance of fibrinogen attachment is that multiple staphylococcal MSCRAMMs are targeted to this critical coagulation factor. Redundancy of function allows for continued pathogenesis despite mutations in the host or pathogen. Furthermore, *S. aureus* virulence factors affect fibrinogen via different mechanisms. Coagulase activates prothrombin to initiate fibrinogen cleavage. ClfA causes clumping of fibrinogen. ClfB, FnbpA and FnbpB are also capable of binding to the multiple chains that make up fibrinogen.

There are two putative mechanisms of virulence in the binding of SdrE2-N2N3 to fibrinogen. First, it has been shown that the addition of SdrE2 inhibits the conversion of fibrinogen to fibrin. The inability of fibrinogen to form clots would present a broader opportunity for *S. aureus* to invade its host. Second, the target site of SdrE2-N2N3 contains an integrin-binding RGD motif. Interfering with the ability of host cells, such as platelets and neutrophils, to adhere to fibrinogen represents a potent potential virulence mechanism for SdrE2. [47]

RESULTS

SdrE1 and SdrE2 have significant differences in affinity for human fibrinogen

As previously published, SdrE2-N2N3 binds to human fibrinogen with high affinity. Given that SdrE2, formerly Bbp, is an allelic variant of SdrE1 with high sequence identity, the ability for SdrE1-N2N3 to bind human fibrinogen was tested in ELISA-type solid-phase assay. SdrE1-N2N3 did show the ability to bind fibrinogen-coated multititer wells; however, it also displayed a lower apparent affinity for human fibrinogen than SdrE2-N2N3 in these experiments (Figure 4-4). Although SdrE1 and SdrE2 share 67% amino acid sequence identity within their N2N3 domains, there was a 5-fold difference in apparent affinity with SdrE1-N2N3 displaying a K_D = 0.5 μM and SdrE2-N2N3 displaying a K_D = 0.1 μM. This indirect detection assay is not a highly accurate way to determine or compare K_D approximations for multiple reasons, chief among them being the different primary antibodies used to detect the respective MSCRAMM binding to the multititer wells.

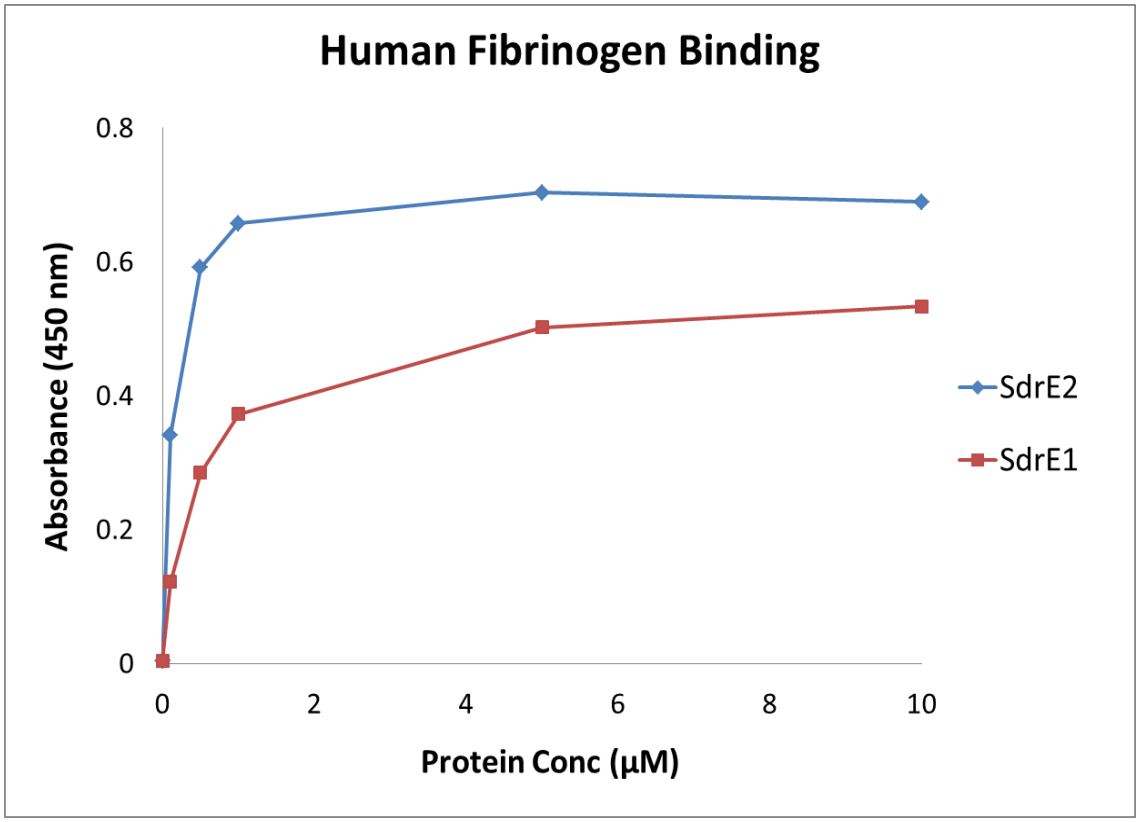


Figure 4-4. SdrE1 and SdrE2 bind Human Fibrinogen.

SdrE1 and SdrE2 bind the same target sequence in the fibrinogen Aα chain

In order to confirm that SdrE1 binds human fibrinogen at the same site as SdrE2, Peptide Inhibition ELISA-type Assays were performed (Figure 4-5). Increasing amounts of the peptide representing the human fibrinogen target sequence were pre-incubated with 0.3 µM SdrE1 or SdrE2. As with SdrE2-N2N3, SdrE1-N2N3 was inhibited from binding to the human fibrinogen coated wells;

however, this inhibition requires a 10-fold higher concentration of peptide to inhibit binding by 80%. SdrE2-N2N3 displayed an IC_{80} = 1.5 μM while SdrE1-N2N3 displayed an IC_{80} = 15 μM. These data give more support to a large difference in affinity for human fibrinogen between SdrE1 and SdrE2.

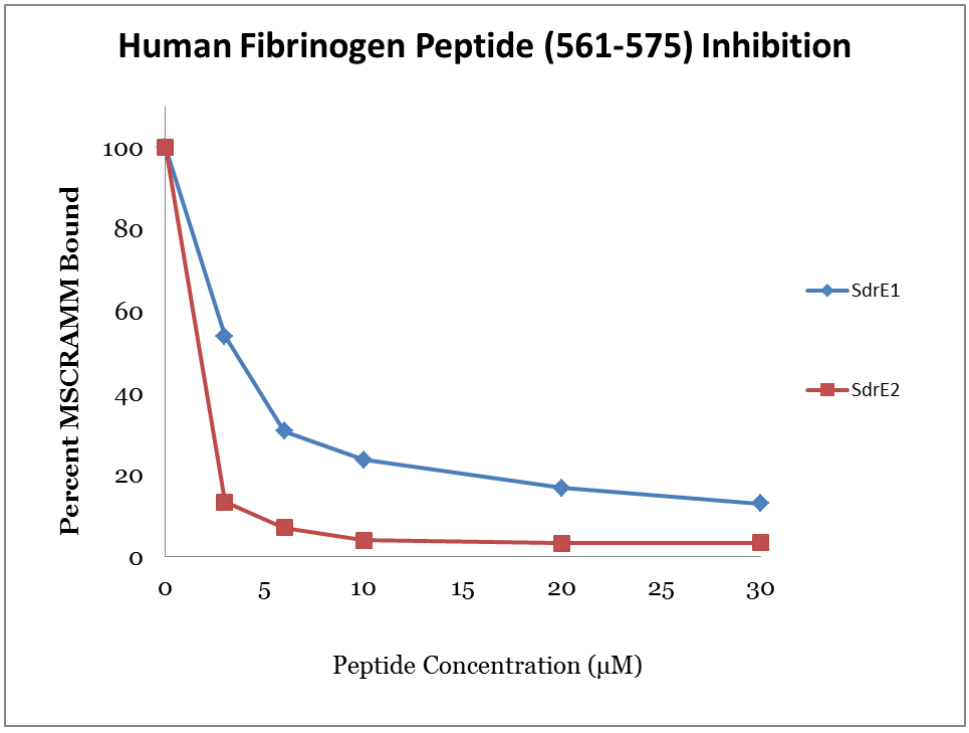
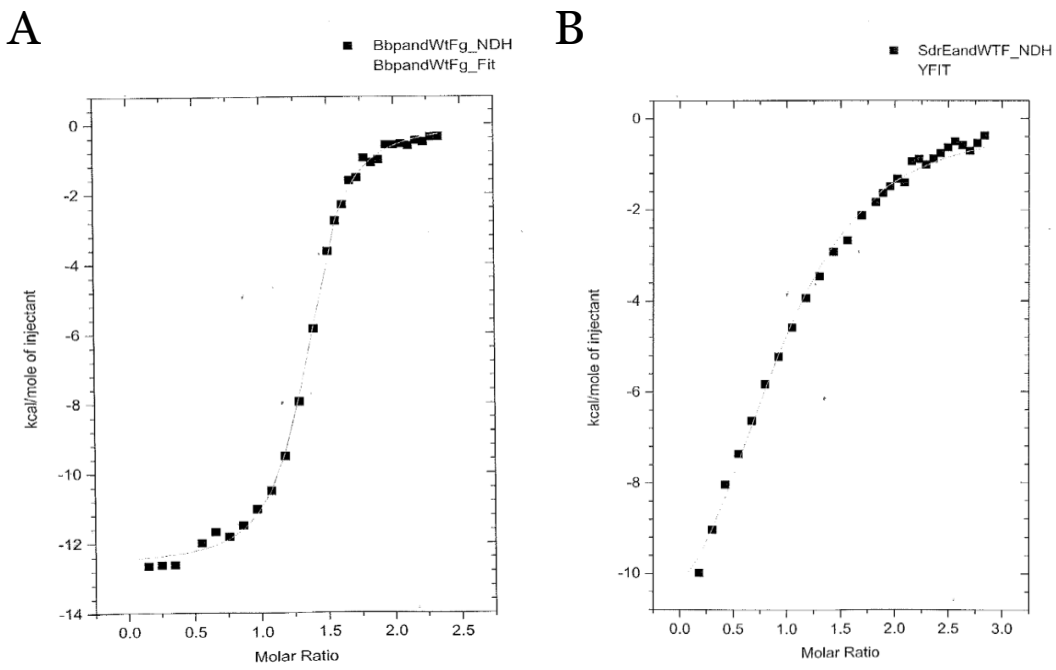


Figure 4-5. Fibrinogen Peptide Inhibition of SdrE binding fibrinogen.

ITC measurements of fibrinogen peptide binding by SdrE1 and SdrE2

In order to more accurately determine the putative differences in affinity, analytical biochemical techniques were used. Isothermal Titration Calorimetry (ITC) is beneficial because it provides an accurate measurement of the thermodynamic parameters of binding and is the best approach to investigate binding of MSCRAMM and ligand peptide in solution. ITC was used to measure the affinity of the two allelic variants for both human fibrinogen and the peptide representing the human fibrinogen target site. Based on previous data, we hypothesized that there would be a significant difference in affinity between SdrE1 and SdrE2 for Fg, with SdrE2-N2n3 displaying a significantly higher affinity than SdrE1-N2N3.

When the data were compared for ITC experiments with the allelic variants binding the previously described human fibrinogen peptide, SdrE2-N2N3 displays a 20-fold greater affinity than SdrE1-N2N3 (Figure 4-6). The differences in affinity for the peptide are traced to differences in the change in entropy between the two interactions. Also, SdrE2-N2N3 having 5 fold greater affinity for full length fibrinogen than SdrE1-N2N3 (data not shown). The difference in affinities for the peptide compared to full length fibrinogen are likely attributed to the greater steric hindrance that the MSCRAMMs must overcome to bind the full length fibrinogen.



	N	K	ΔH	ΔS	Kd (calc)
SdrE2	1.335	3.818 E 6	-1.259 E 4	-11.43	0.261 µM
SdrE1	0.94	2.315 E 5	-1.324 E 4	-19.2	4.3 µM

Figure 4-6. ITC Measurement of SdrE1 and SdrE2 binding Fibrinogen Peptide

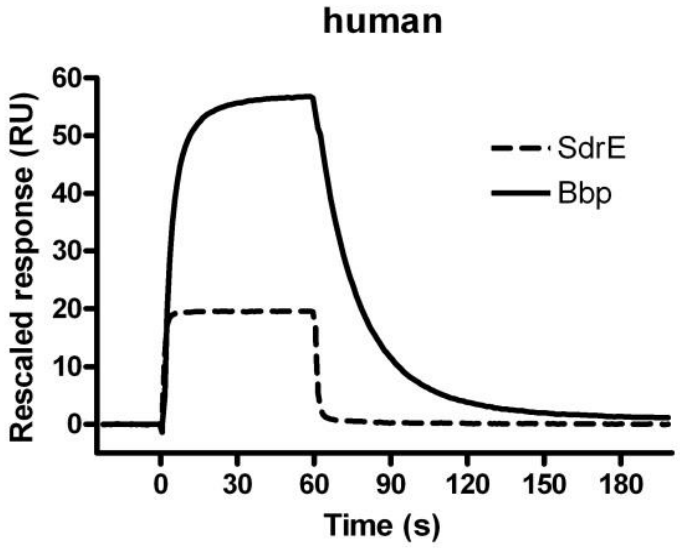
(A) SdrE2-N2N3 binds the peptide representing the target site in human fibrinogen with much greater affinity than (B) SdrE1-N2N3 binds the same peptide. The table below shows the binding parameters elucidated from these experiments.

When ClfA-N2N3 binding of full length fibrinogen and a peptide representing the binding site in fibrinogen are examined in ITC experiments, ClfA-N2N3 was observed to bind to full length fibrinogen better than the peptide. Further experimentation confirmed the underlying reason for this observation is a secondary fibrinogen binding site outside of the ligand binding trench on ClfA-N2N3. Given that both SdrE allelic variants are observed to bind to the peptide with a greater affinity than to the full length fibrinogen, as well as the N numbers displayed in these experiments, the data suggest that the previously elucidated target site is the only major binding site of SdrE in fibrinogen.[34]

SPR measurements of human fibrinogen binding by SdrE1 and SdrE2

Surface Plasmon Resonance was used to investigate the interactions between SdrE1 and SdrE2 and human fibrinogen. It was previously reported that SdrE2 binds human fibrinogen coated chips with a K_D = 0.54 μM ± 0.07. When repeated, we calculate a K_D = 1 μM for SdrE2-N2N3 to human fibrinogen. SdrE1-N2N3 has a K_D = 46 μM (Figure 4-7). The affinity of SdrE2-N2N3 for full length fibrinogen is similar in both ITC and SPR experiments. However, there is a 7.5-fold difference between the affinity constants of SdrE1-N2N3 as measured in the two techniques. Both experiments were repeated multiple times, leaving us to hypothesize that the differences are due to changes in the display of the binding

site between coated and soluble fibrinogen. While the magnitude of the difference in binding varies by technique, the binding profiles remains consistent across each set of experiments. SdrE1-N2N3 displays a lower affinity, broad species specificity fibrinogen binding profile while SdrE2-N2N3 displays a high affinity profile that is highly specific to human and cow fibrinogen.



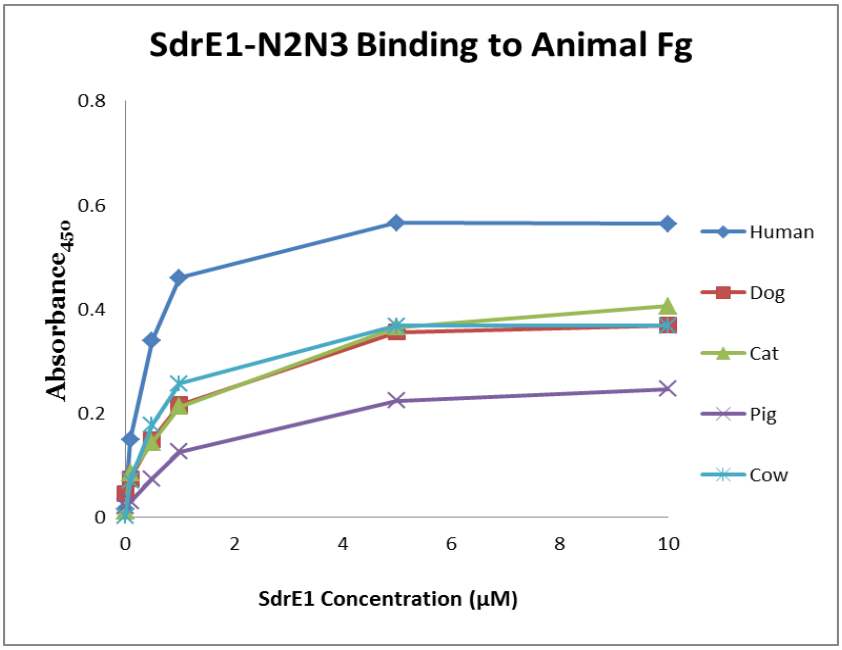
MSCRAMM / Species of Fg immobilized	SdrE		Bbp	
	K_D (µM)	*N*	K_D (µM)	*N*
Human	46	*1.50*	0.99	*0.62*

Figure 4-7. SPR measurement of the SdrE-Fibrinogen Interaction.

SdrE2 has high specificity for human fibrinogen while SdrE1 has broad species specificity

As previously published, SdrE2 binds specifically to human fibrinogen with high affinity, but does not interact strongly with fibrinogen from other animals (Figure 4-8). The high specificity of SdrE2 for human fibrinogen is especially intriguing given the high sequence similarity between the target human fibrinogen sequence and the corresponding sequences from some animals such as dogs. In all fibrinogen samples from non-human species tested here, there is an insertion of a valine (Table 4-1). It appears likely that this valine is responsible for changes in binding.

While SdrE2-N2N3 showed a highly specific binding profile, SdrE1-N2N3 showed a significantly different profile in ELISA-type assay featuring microtiter wells coated with fibrinogen from different species. SdrE1-N2N3 is able to bind with varying affinity to fibrinogen from multiple animals. The differing binding profiles are particularly noteworthy given the 67% amino acid sequence identity shared between the N2N3 domains of these proteins.



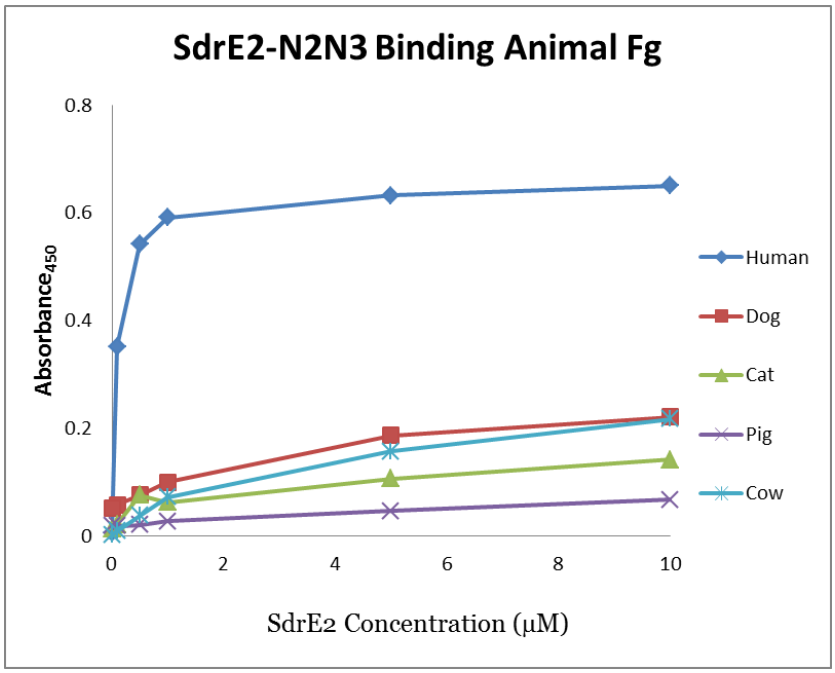


Figure 4-8. SdrE1 and SdrE2 binding to animal fibrinogen

Table 4-1. Corresponding Animal Fibrinogen Sequences.

Human	SKQF-TSSTSYNRGDS
Dog	SKQFVTSSTTYNRGDS
Cat	SKQLVATSKTYNRGDS
Pig	SKQTIT--KTINREGR
Cow	SKQFVSSSTTVNRGGS

SPR measurements of animal fibrinogen binding by SdrE1 and SdrE2

SPR experimentation was used to confirm the findings of the animal fibrinogen ELISA-type assays. The ability of SdrE1-N2N3 and SdrE2-N2N3 to bind to chips coated with fibrinogen from humans, cows, pigs, dogs, cats or cows was measured to determine the kinetic parameters of these interactions, including K_D (Figure 4-9). In addition to its high affinity interaction with human fibrinogen, SdrE2 was able to interact with cow fibrinogen, but with a 20-fold lower affinity compared to the SdrE2-human fibrinogen interaction. In contrast, SdrE1 was able to interact with fibrinogen from all animals tested (Table 4-2) with affinities ranging from 1.5-fold to 4-fold lower than the affinity of SdrE2 for human fibrinogen.

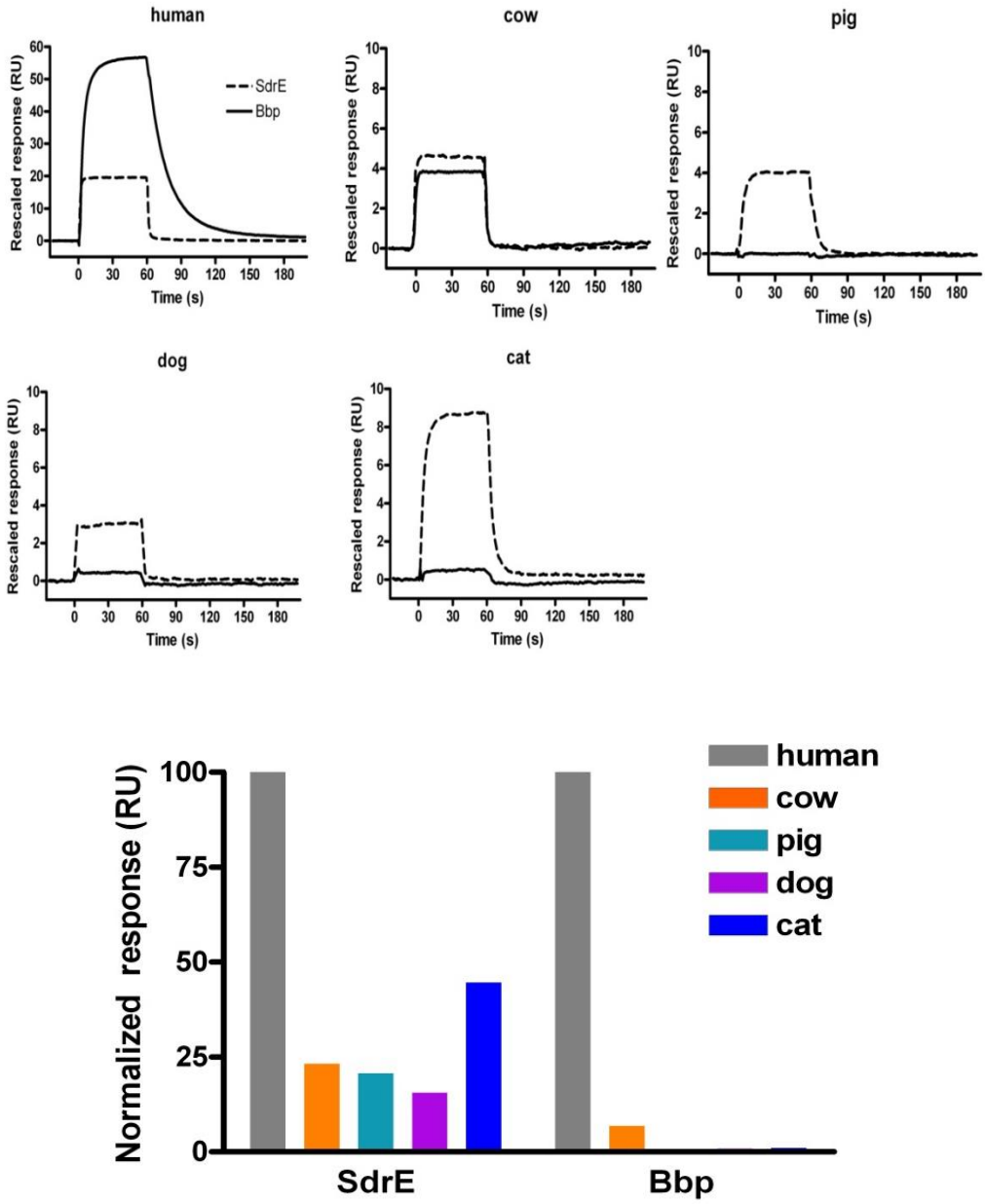


Figure 4-9. SPR measurements of animal fibrinogen binding by SdrE1 or Bbp/SdrE2.

Table 4-1. Binding Parameters calculated from SPR Experiments.

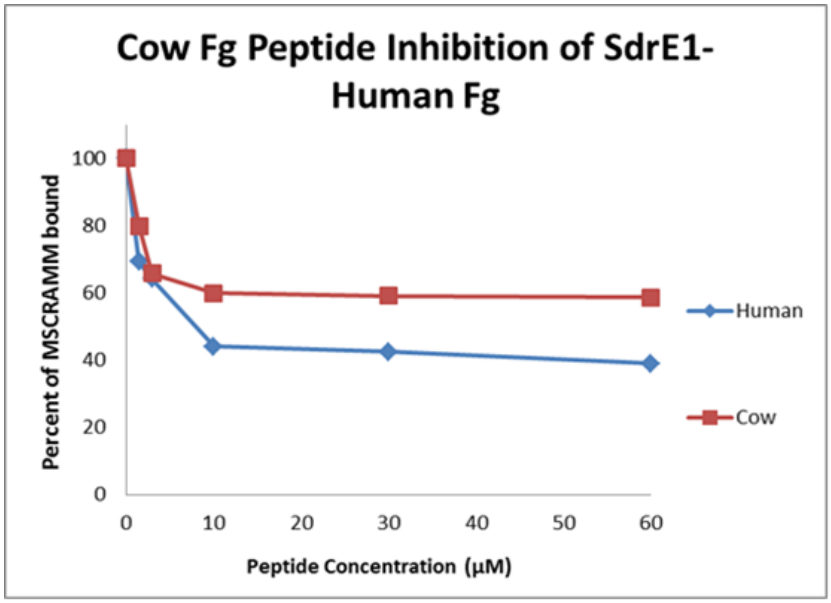
MSCRAMM / Species of Fg immobilized	SdrE		Bbp	
	K_D (μM)	*N*	K_D (μM)	*N*
Human	46	*1.50*	0.45 7.10	*0.36* *0.43*
Cow	130	*0.98*	20	*0.22*
Pig	140	*1.01*	ND	*ND*
Dog	143	*0.83*	ND	*ND*
Cat	67	*1.04*	ND	*ND*

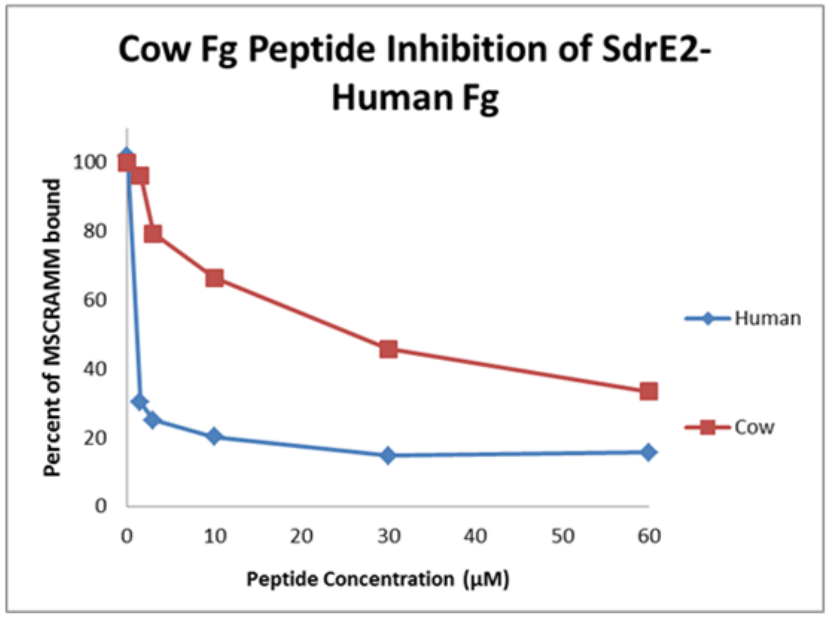
SdrE1 and SdrE2 bind to the same site in cow and human fibrinogen

While the ELISA-type assay was not sensitive enough to detect binding of SdrE2-N2N3 to cow fibrinogen, there was a measurable interaction between these two proteins in SPR experiments. When comparing the target sequence in human fibrinogen to the corresponding sequence in cow fibrinogen, there are only four changes. Based on the similarities, we hypothesized that the binding of cow fibrinogen by SdrE1 and SdrE2 was targeted to this region. To test this, a Peptide Inhibition ELISA-type assay was performed wherein 0.5 μM SdrE1 or SdrE2 was pre-incubated with an increasing amount of human or cow fibrinogen

peptide (1.5-60 μM), then added to wells coated with human fibrinogen (Figure 4-10). The cow fibrinogen peptide was able to inhibit the binding of SdrE1 and SdrE2 to human fibrinogen coated wells, but to a lesser degree than the human fibrinogen peptide. This provides evidence that the target site in cow fibrinogen corresponds to the target site in human fibrinogen. It also corroborates the previous data that these allelic variants have a higher affinity for human fibrinogen than animal fibrinogen.

The Peptide Inhibition ELISA-type assay technique was also performed using peptides representing the amino acid sequences from other species that correspond to the human fibrinogen target sequence (Figure 4-11). SdrE2-N2N3, which does not bind to fibrinogen from species other than human or bovine, was not inhibited by peptides representing the target site in fibrinogen from dogs, cats or pigs. SdrE1-N2N3, which shows a more promiscuous fibrinogen binding profile in ELISA-type assays, was inhibited at rates ranging from 20% to 60% by the different target peptides. Given that human fibrinogen is still the species for which SdrE1-N2N3 shows the highest affinity, it is not surprising that the animal fibrinogen peptides were unable to completely inhibit binding even at high concentrations.

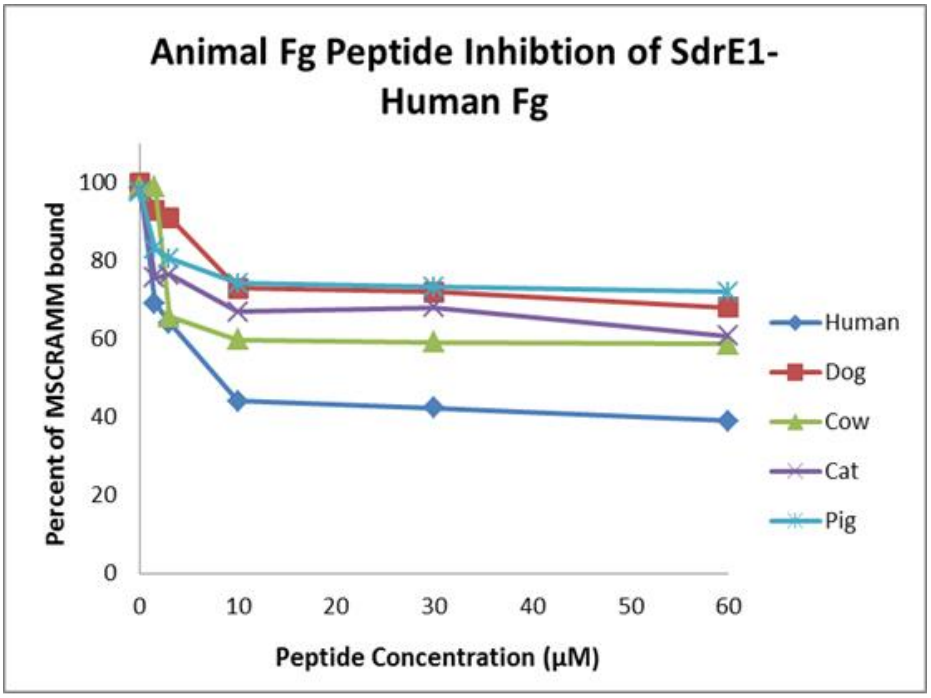




Human SKQF-TSSTSYNRGDS

Cow SKQFVSSSTTVNRGGS

Figure 4-10. Cow Fibrinogen Peptide Inhibition.



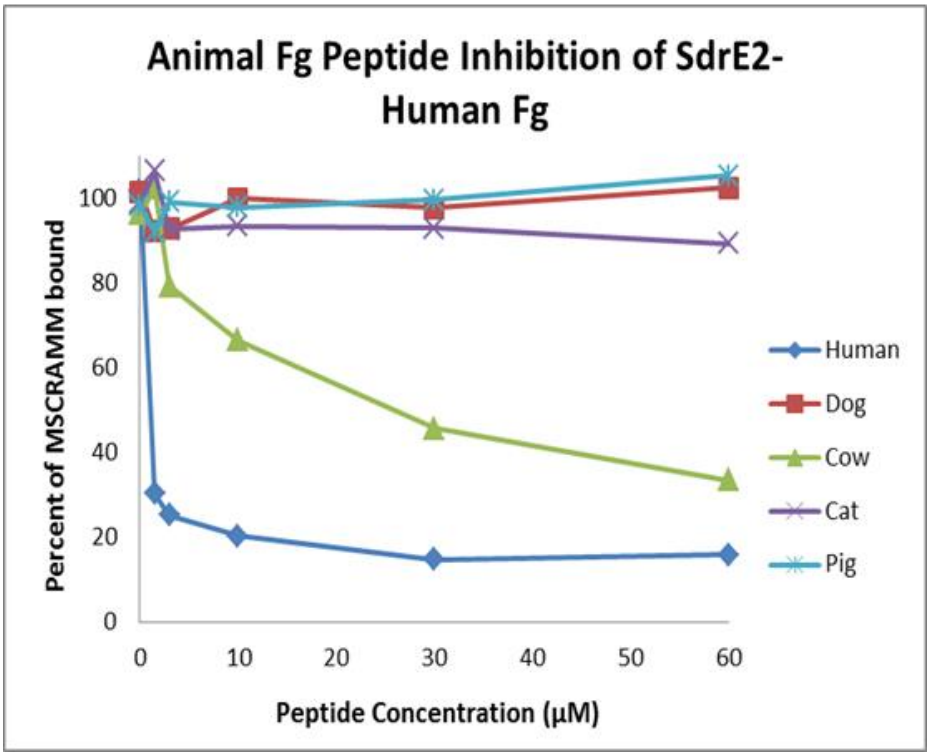


Figure 4-11. Animal Fibrinogen Peptide Inhibition.

ITC measurements of SdrE1 and SdrE2 binding cow fibrinogen peptide

ITC was used to measure the affinity of the two allelic variants for the cow fibrinogen peptide. In SPR experiments using full length cow fibrinogen, a similar maximal binding was observed, but a significantly different K_D was calculated between the two. In ITC experiments using a soluble peptide representing the fibrinogen binding site, SdrE1-N2N3 and SdrE2-N2N3 displayed a similar affinity.

DISCUSSION

While SdrE1 and SdrE2 are 67% identical within the ligand binding N2N3 domains, they have significantly different biochemical profiles with regards to fibrinogen binding. SdrE2-N2N3 binds with high affinity and high specificity to human fibrinogen, while SdrE1-N2N3 binds with lower affinity to human fibrinogen but displays broad species specificity in fibrinogen binding.

Interestingly, these results give a biochemical rationale for the observed gene frequencies of SdrE1 and SdrE2 in human and animal staphylococcal isolates (Chapter 2). s*drE1* is present at higher frequencies in animal isolates and SdrE1-N2N3 able to interact with fibrinogen from these animals. s*drE2* is present at higher frequencies in human isolates as compared to its presence in animal isolates and SdrE2-N2N3 shows little affinity for fibrinogen from non-

human species. An outlier to this framework is the cow. The otherwise human-specific SdrE2-N2N3 did show an affinity for cow fibrinogen as well as a sequence in cow fibrinogen that corresponds to the target sequence in human fibrinogen. It is interesting to note that out of the staphylococcal isolates gathered from animals that had the *sdrE2* gene, 75% were gathered from bovine sources.

From these data, we propose that SdrE2 is an example of adaptation by *S. aureus* to the human host in which accumulated mutations resulted in improved binding to human fibrinogen and a loss of binding to fibrinogen from most other species. The high affinity demonstrated for fibrinogen by multiple analytical biochemical techniques suggests that fibrinogen binding is the main role of SdrE.

It is important to note that there are relatively few changes between *sdrE1* and *sdrE2* that resulted in this dramatic change in phenotype. However, within the N2N3 domains, there are still 115 differences in amino acid sequence. It is likely that not all of these changes are required for the change in phenotype.

CHAPTER V

STRUCTURAL BASIS FOR DIFFERENCES IN FIBRINOGEN BINDING PROFILE OF SDRE1 AND SDRE2

INTRODUCTION

The fibrinogen binding profile of allelic variants SdrE1-N2N3 and SdrE2-N2N3 provide a rationale for the observed epidemiology of their respective genes. The magnitude of the differences in binding, as measured with multiple techniques, was larger than one would expect in light of the high degree of similarity between the amino acid sequences of the variants. There is 87% sequence identity between the genes, including 95% identity outside of the N2N3 domains and 67% identity within the N2N3 ligand binding domains.

We hypothesized that mutational analysis, based on structural data gathered from the crystal structures of apo- and bound SdrE1-N2N3 and SdrE2-N2N3, would provide insight into the specific differences between SdrE1 and SdrE2 that are responsible for the differences in fibrinogen binding. While the residues that are different between the variants are known and could be modeled based on the solved crystal structures of other similar MSCRAMMs, true insight into the molecular nature of the SdrE1/2-Fibrinogen interaction would require experimental data in the form of crystal structures of apo-SdrE1-N2N3, apo-

SdrE2-N2N3, SdrE1-N2N3-FgPeptide co-crystal, and SdrE2-N2N3-FgPeptide co-crystal.

Towards this goal, recombinant SdrE1-N2N3 and SdrE2-N2N3 were expressed and purified. These purified proteins were used in collaboration with Dr. Ganesh Vannakambadi who is responsible for crystallization and structure solving. Initially, the apo-SdrE2-N2N3 and SdrE2-N2N3-FgPeptide co-crystal were solved. Mutational analysis of the target site in fibrinogen and SdrE2-N2N3 were carried out based on these structures. Recently, the structures of apo-SdrE1-N2N3 and SdrE1-N2N3-FgPeptide were also solved. To support the notion that the differences in fibrinogen binding by SdrE allelic variants are driven by a small set of SdrE residues, we generated SdrE1/SdrE2 chimeric constructs. We hypothesized that the binding profile of one allelic variant could be mimicked by the other variant with relatively few changes to the amino acid sequence. Based on solved co-crystal structures, we predicted that a chimeric SdrE2 mutant would display the binding profile of SdrE1 – lower affinity for human fibrinogen and broad specificity for a variety of fibrinogen species - by changing a minimal number of critical SdrE2 to their respective SdrE1 counterparts. Similarly, with the solved SdrE1-N2N3-Fg structure, we hypothesize that a chimeric SdrE1 molecule could be converted to a high affinity and high specificity fibrinogen binding profile found in wild-type SdrE2.

RESULTS

Crystal structures of apo-SdrE2-N2N3 and SdrE2-N2N3-Fg peptide co-crystal

Dr. Ganesh Vannakambadi solved the structures for apo-SdrE2-N2N3 and the SdrE2-N2N3-Fg Peptide co-crystal (Figure 5-1). These data show that SdrE2 binds to fibrinogen via the Dock, Lock and Latch model in a similar manner to other MSCRAMMS. In this model, fibrinogen encounters SdrE2 in the open conformation and enters the ligand binding trench formed by the N2 and N3 domains of SdrE2. Subsequently, the Lock domain closes the conformation by covering the ligand-filled trench. Finally, the Latch domain inserts into the N2 domain by β-strand complementation. As seen in Figure 5-1, the conformational changes between the open, empty state and the closed, bound state are mainly found in the flexible Lock and Latch domains.

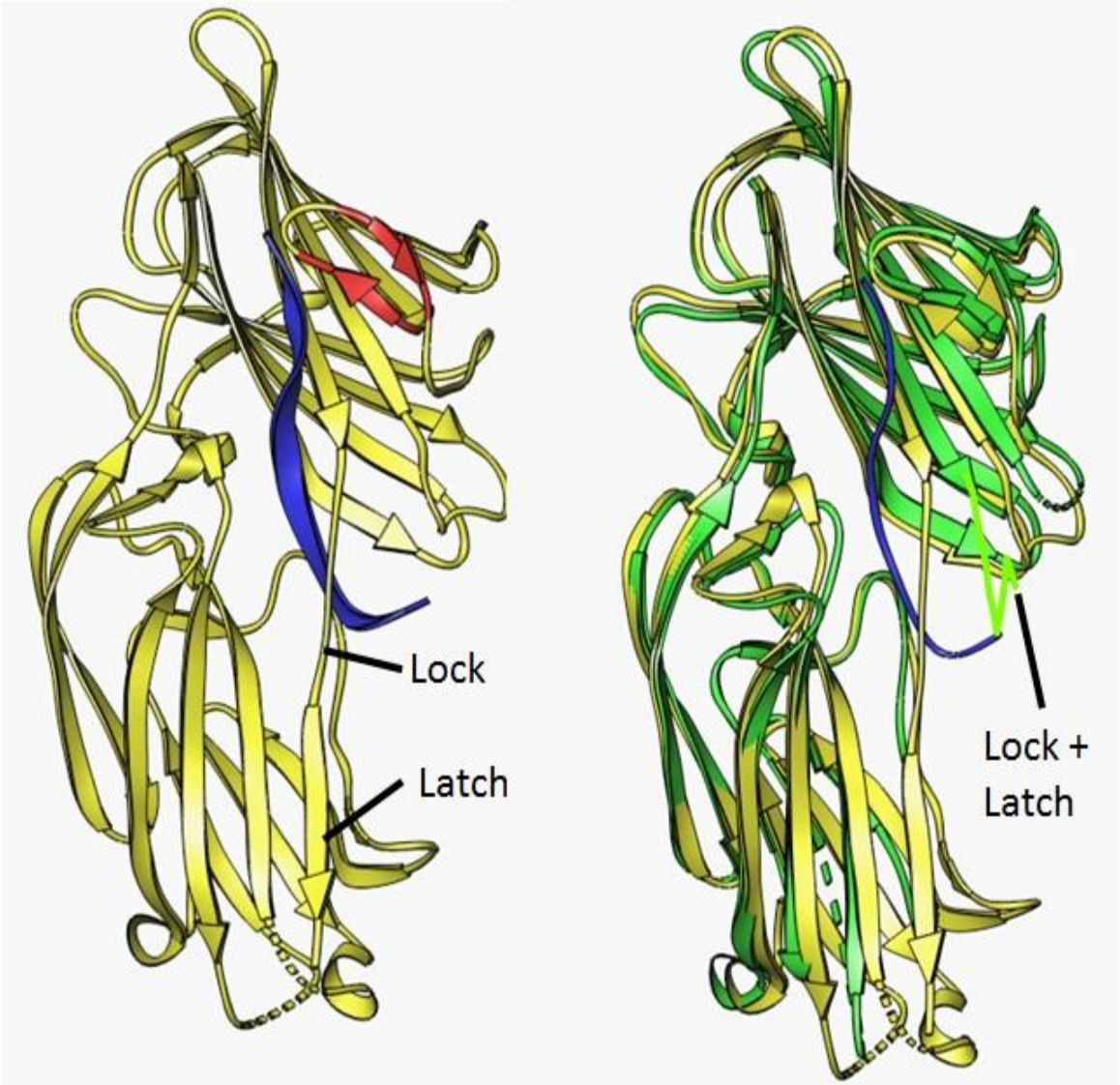


Figure 5-1. Apo-SdrE2-N2N3 and SdrE2-N2N3-Fg Peptide Co-Crystal.
Panel 1 shows SdrE2-N2N3 (yellow) structure when bound to the human fibrinogen peptide (blue) that is representing the target site. New, additional β-hairpin shown in red. Panel 2 shows an overlay of this structure with the structure of Apo-SdrE2-N2N3 (green).

While SdrE2-N2N3 binds Fg via the Dock, Lock and Latch model, there are a few important differences from previous examples of Dock, Lock and Latch binding. There is a small β-hairpin (Fig 5-1, red) that intrudes on the ligand binding trench and is not found in other MSCRAMMs. The β-hairpin is seen in both the open and closed conformations of SdrE2-N2N3. This creates additional contact points with the ligand and causes the ligand to form a heretofore unseen twist within the trench (blue). Also, densities were only seen for the residues 561-573 of the peptide.

Mutational analysis of the target sequence

Previously, the specific sequence in fibrinogen to which SdrE2-N2N3 binds was isolated via binding assays with truncation mutants of the fibrinogen Aα chain. Through these experiments, residues 561-575 were identified as the site of binding. A 15 amino acid peptide representing this binding site was used to confirm this finding in Peptide Inhibition ELISA-type assays and ITC.

SdrE2-N2N3 does not bind to fibrinogen from non-human species with the exception of cow fibrinogen. This is particularly noteworthy in light of the high degree of sequence similarity between the target sequence in human fibrinogen and the corresponding sequences from fibrinogen from other species. In order to understand the structural basis for this, peptides were synthesized that contained point mutations in the wild type human fibrinogen sequence. The

mutations were designed by Dr. Ganesh Vannakambadi based on the solved crystal structure of SdrE2-N2N3 in complex with the human fibrinogen peptide. A Peptide Inhibition ELISA-type assay was used to assess the relative affinities of the mutant peptides for SdrE2-N2N3. It was observed within the structure that residues 573-575 (Gly-Asp-Ser) did not appear to be interacting with any of the residues within the MSCRAMM. It was hypothesized that a peptide lacking these residues would display unaltered affinity for SdrE2-N2N3 in comparison to the original peptide. A peptide representing residues Fg Aα 559-572 ("shift") was ordered and tested in the Peptide Inhibition ELISA-type assay (Figure 5-2). The resulting inhibition curve was not statistically different from the original (WT) fibrinogen peptide. This suggests that GDS residues are not important for binding, confirming the hypothesis.

That SdrE2-N2N3 binding of fibrinogen was effected by minor changes in sequence was of particular interest in these studies. The most notable change from human fibrinogen to fibrinogen from other species is the insertion of a valine at position 575. This insertion is shared by fibrinogen from all species investigated. The insertion of a valine potentially alters the register of the SdrE2-Fg interaction.

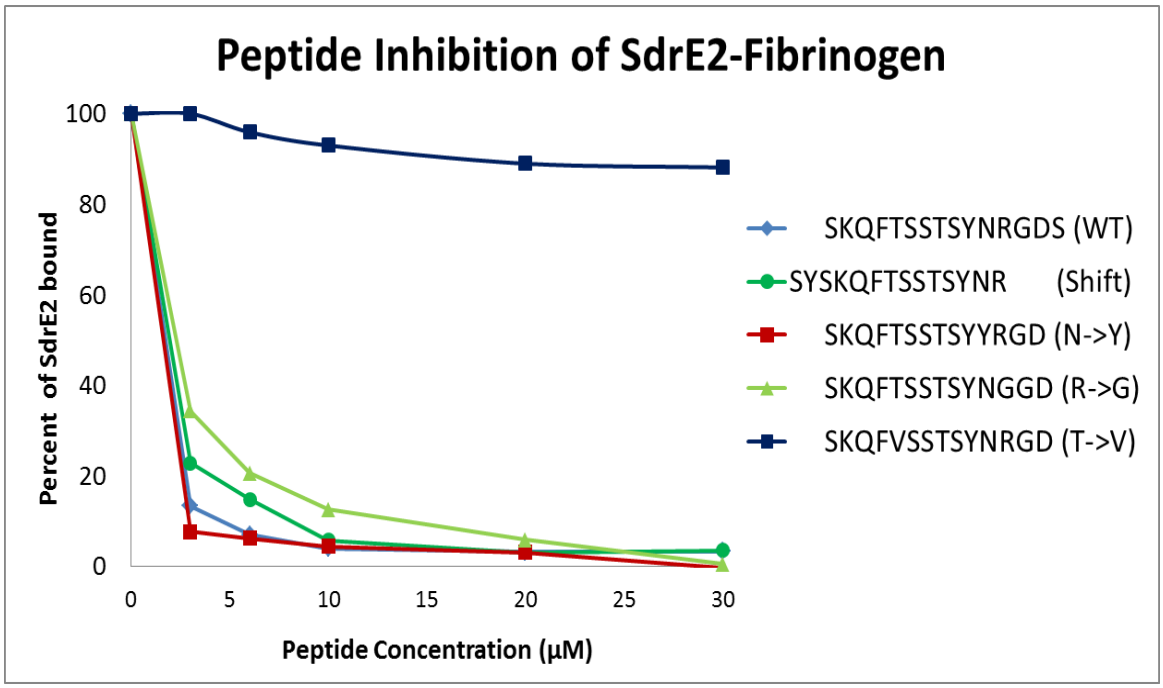


Figure 5-2. Mutational Analysis of the Fibrinogen Target Site.
Peptide inhibition of the ability of SdrE2-N2N3 to bind to human fibrinogen coated wells was performed as an indirect measure of binding affinity for these mutated peptides.

In order to determine which shifted residues were responsible for a reduction in binding, three separate point mutations were made within the original fibrinogen peptide: T565V, N571Y, and R572G. Each mutation replaces a putatively important wild type residue with the residue that would occur at that position with a valine insertion at position 575. Without the solved structures, many more mutations would have been necessary to determine critical residues. The structural data allowed us to narrow our focus.

N571Y showed no difference in binding from the wild type peptide. R572G consistently showed a moderate decrease in binding, suggesting that interaction with this residue is of some importance for SdrE2-N2N3. The largest difference in binding occurred with the T565V mutation, where inhibition was not seen at the highest concentration of peptide. This strongly suggests that SdrE2-N2N3 does not interact strongly with the peptide because it is instead binding to the human fibrinogen-coated wells. In a follow-up assay, partial inhibition was eventually seen at 1000 times the concentration of MSCRAMM (Figure 5-3); this represents a >100-fold lower affinity.

Mutational analysis of SdrE2-N2N3

The solved structure of SdrE2-N2N3-Fg Peptide co-crystal at high resolution reveals the residues in the MSCRAMM and peptide whose close proximity strongly suggest an interaction. In order to confirm these interactions, mutational analysis of SdrE2-N2N3 was performed. Putative residues integral for binding were determined by Dr. Ganesh Vannakambadi. Constructs containing the point mutations listed in Table 5-1 were made using site-directed mutagenesis. Constructs confirmed to have the desired mutations by sequencing were transformed and expressed in E. coli, purified through a 2-step affinity chromatography process, and then finally tested in ELISA-type assay for binding to fibrinogen-coated wells.

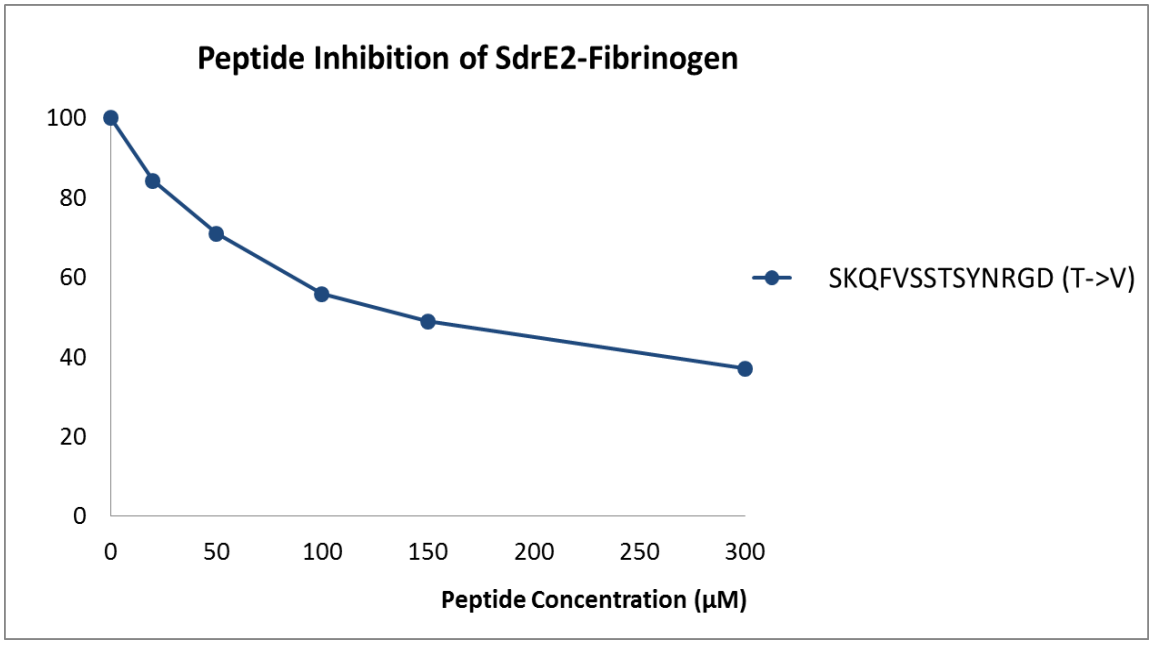


Figure 5-3. Peptide Inhibition of SdrE2 - Fibrinogen with T585V Mutant.

Table 5-1. SdrE2 Mutational Analysis

Name	Purpose
SdrE2-Y475A	Confirm Structure
SdrE2-A469D	Confirm Structure
SdrE2-D588A	Confirm Structure
SdrE2-D480A	Chimeric Mutation
SdrE2-Y475K	Chimeric Mutation
SdrE2-I335L	Chimeric Mutation

While the SdrE2-Y575A mutation did not result in any change in phenotype, the SdrE2-A469D point mutation did show a partial reduction in binding. The most significant change in binding was seen with the SdrE2-D588A mutation where there was a total loss of fibrinogen binding (Figure 5-4). This residue occurs in the lock region (Figure 5-5) and appears to interact with the Q563 residue of the fibrinogen peptide; the Q563 residue has shown to play an important role in binding (Chapter 6).

Interestingly, the D588A mutation results in the same phenotype as seen in the SdrE2-ΔLock and SdrE2-ΔLockLatch deletion mutants. These deletion mutant constructs were made by Vanessa Vazquez to help elucidate the mechanism of SdrE2 binding to fibrinogen. They were purified according to the same protocol as the recombinant wild type N2N3 proteins. Subsequently, these purified proteins were tested in the same ELISA-type assays as previously described.

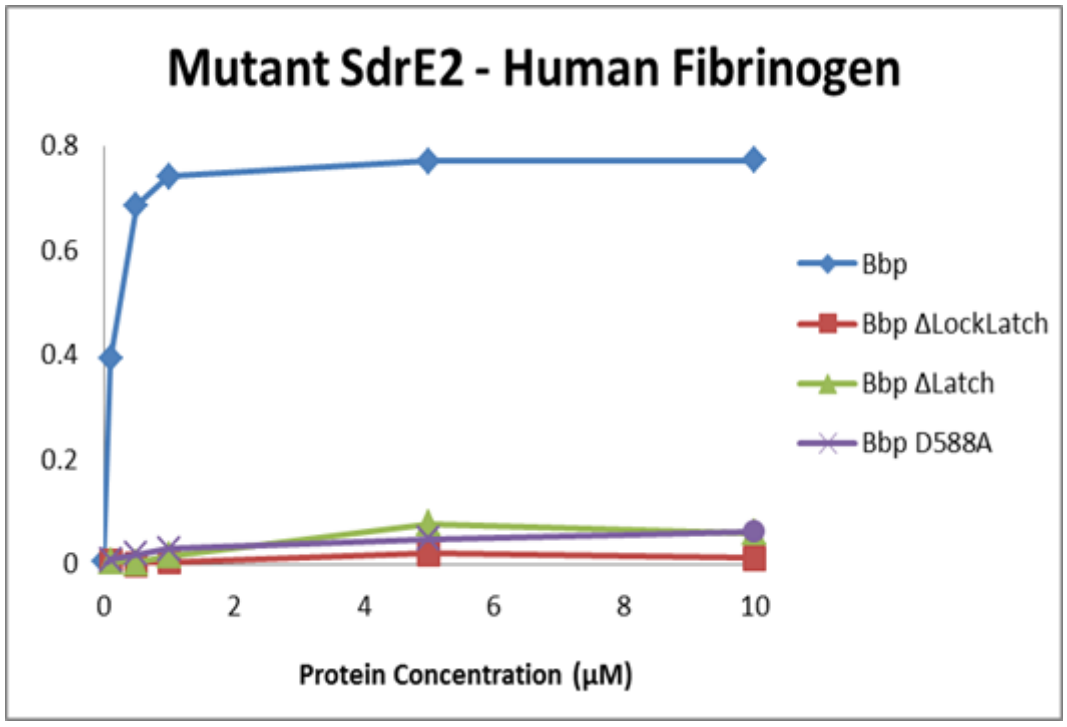


Figure 5-4. Loss of function mutations in SdrE2.

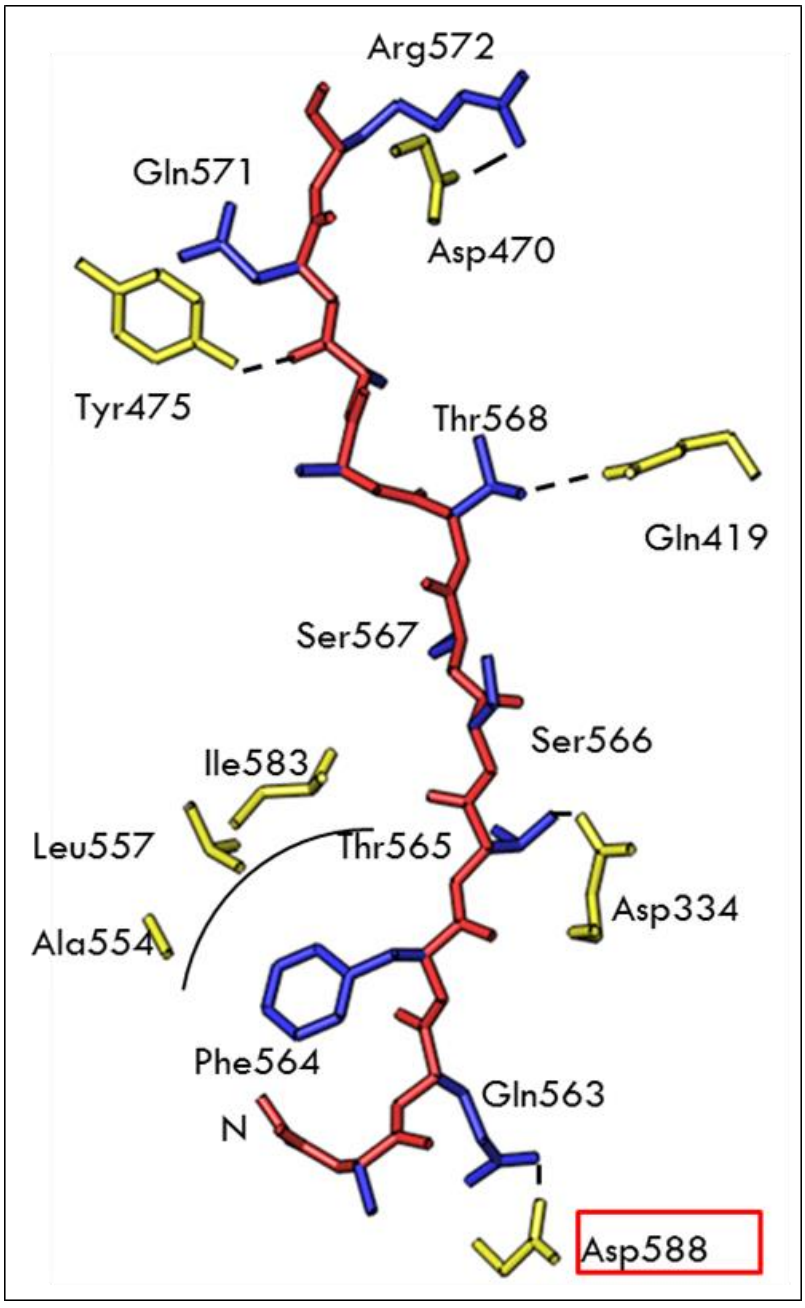


Figure 5-5. Ligand Binding Trench of SdrE2-N2N3.
Human fibrinogen peptide (backbone in red, side chains in blue) shown with the SdrE2 side chains (yellow) that are the shortest distance from the peptide and are the most likely to interact. SdrE2-D588 noted with red box; this residue was mutated to an alanine and showed total abrogation of binding.

No single point mutation creates a chimeric SdrE2 that has the SdrE1 phenotype

Using the solved structures of apo-SdrE2-N2N3 and SdrE2-N2N3-Fg Peptide co-crystal, SdrE2-N2N3 residues were selected that both have a putative role in fibrinogen binding and are different from the corresponding residues in SdrE1-N2N3. Point mutations were made in the wild type SdrE2-N2N3-pQE30 construct using site directed mutagenesis that replaced the SdrE2 residues with the corresponding SdrE1 residue. These mutant constructs were then expressed, purified and tested in ELISA-type assay for their dose-dependent binding to microtiter wells coated with fibrinogen from various species. Representative assays are shown in Figure 5-6. None of the mutants showed a phenotype resembling SdrE1-N2N3 or the repeatable ability to bind to fibrinogen from animals in these assays.

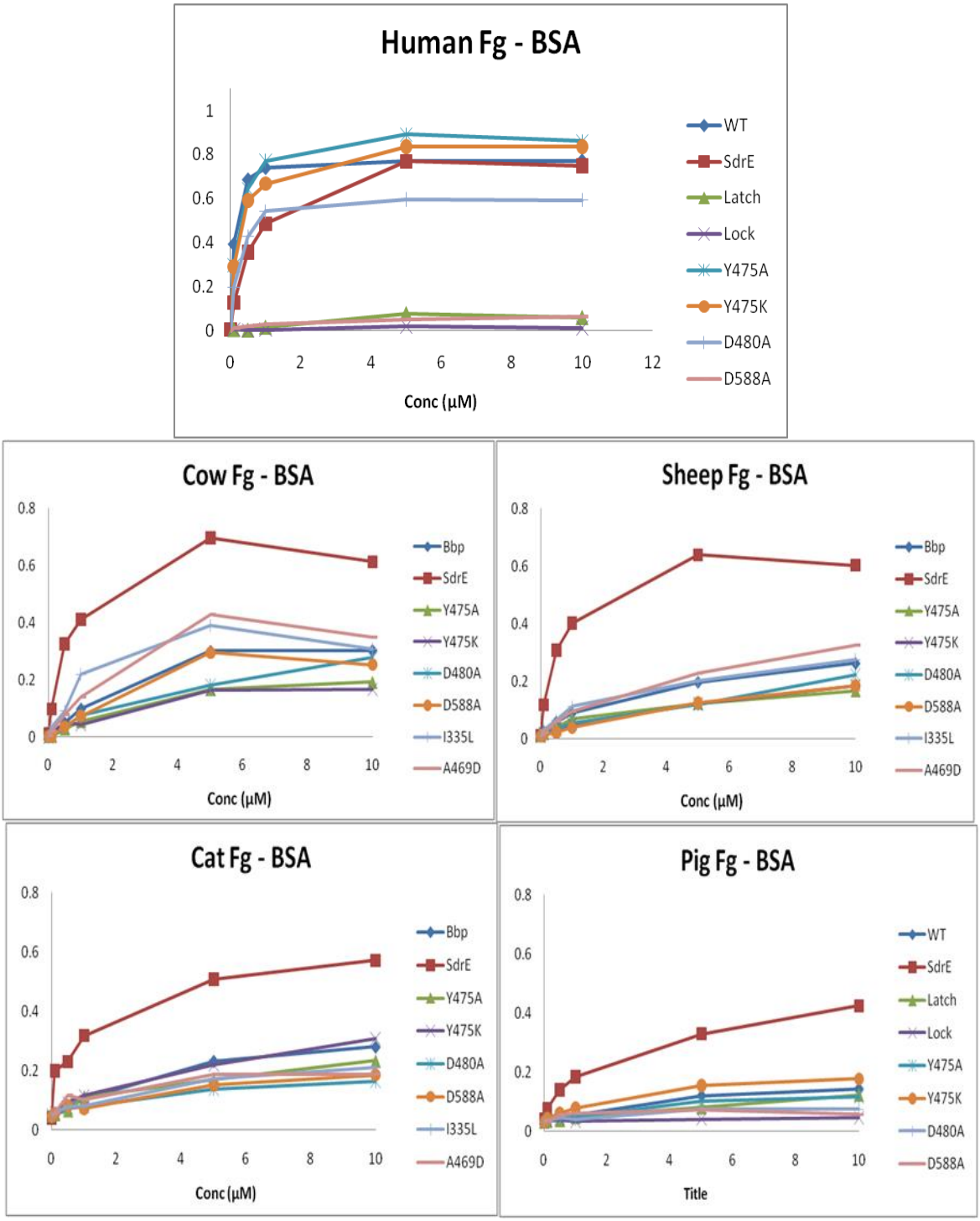


Figure 5-6. Animal Fibrinogen Binding of SdrE2 Chimeric Point Mutants.

While there were some changes in binding to human fibrinogen, none of the listed point mutants were able to replicate the promiscuous binding profile of SdrE1 (red)

SdrE2-LockChimera confers SdrE1 phenotype on SdrE2

Building on the observation that the D588 residue is integral for SdrE2-N2N3 binding of Fg, the residues surrounding D588 were analyzed. D588 is located in the flexible Lock region of SdrE-N2N3 that closes the conformation of the MSCRAMM by covering the trench and interacting with the target sequence. D588 appears to interact with Q563 in the Fg Aα chain.

When the lock regions of SdrE1 and SdrE2 are aligned, there is a significantly higher frequency of variations within the Lock domain than within the two allelic variants overall (Figure 5-7). A hypothesis was formulated that replacing the SdrE2-Lock region with the SdrE1-Lock region would partially change the phenotype (as measured in ELISA-type assay) of SdrE2-N2N3 to that of SdrE1-N2N3. The SdrE2-LockChimera construct was made by using the Overlap Extension PCR technique and then expressed, purified and screened in ELISA-type assay for its ability to bind to microtiter wells coated with fibrinogen from multiple species.

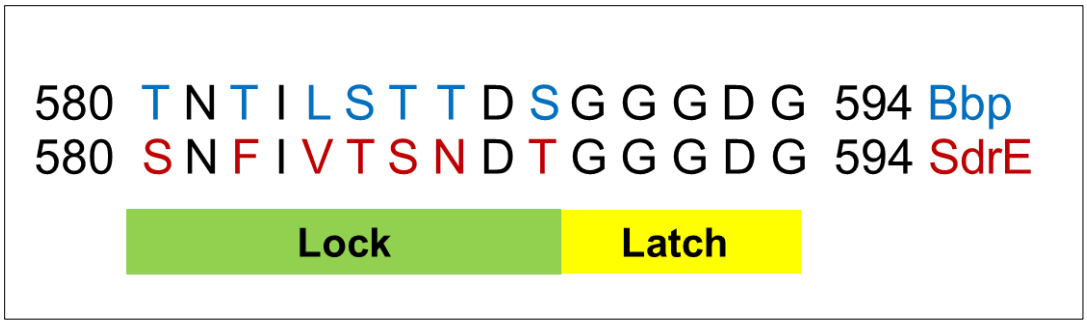


Figure 5-7. SdrE Lock-Latch region alignment.

SdrE2-LockChimera is constituted of 97% SdrE2 residues and 3% SdrE1 residues and is detected in this assay with primary antibody specific to SdrE2-N2N3. Despite this, SdrE2-LockChimera shows a binding profile more similar to the lower affinity, broad specificity binding of SdrE1-N2N3 than SdrE2-N2N3 (Figure 5-8). The affinity for human fibrinogen appears to be reduced with a 2-5 fold change in apparent KD when comparing the SdrE2-N2N3 and SdrE-LockChimera. However, the chimeric mutant still appears to have a greater affinity for human fibrinogen than SdrE1-N2N3.

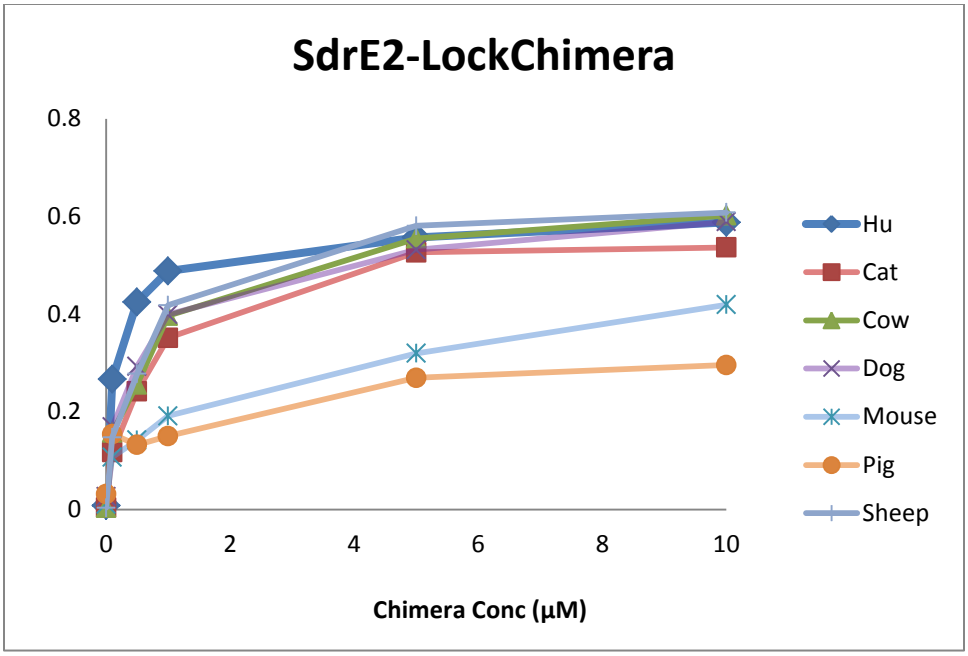


Figure 5-8. SdrE2-Lock Chimera binds fibrinogen from animals.

The apparent reduction in affinity was tested in ITC using the human fibrinogen peptide mentioned previously (FgAα561-575). SdrE2-LockChimera showed a KD = 0.86 µM, a 3.5-fold reduction in affinity for this peptide when compared to SdrE2-N2N3. However, when compared to SdrE1-N2N3, the SdrE2-Lock Chimera still has a 5-fold stronger affinity (Figure 5-9). This suggests that while the lock region plays a very important role in binding, there are other residues that contribute to the difference in binding profile between SdrE1-N2N3 and SdrE2-N2N3.

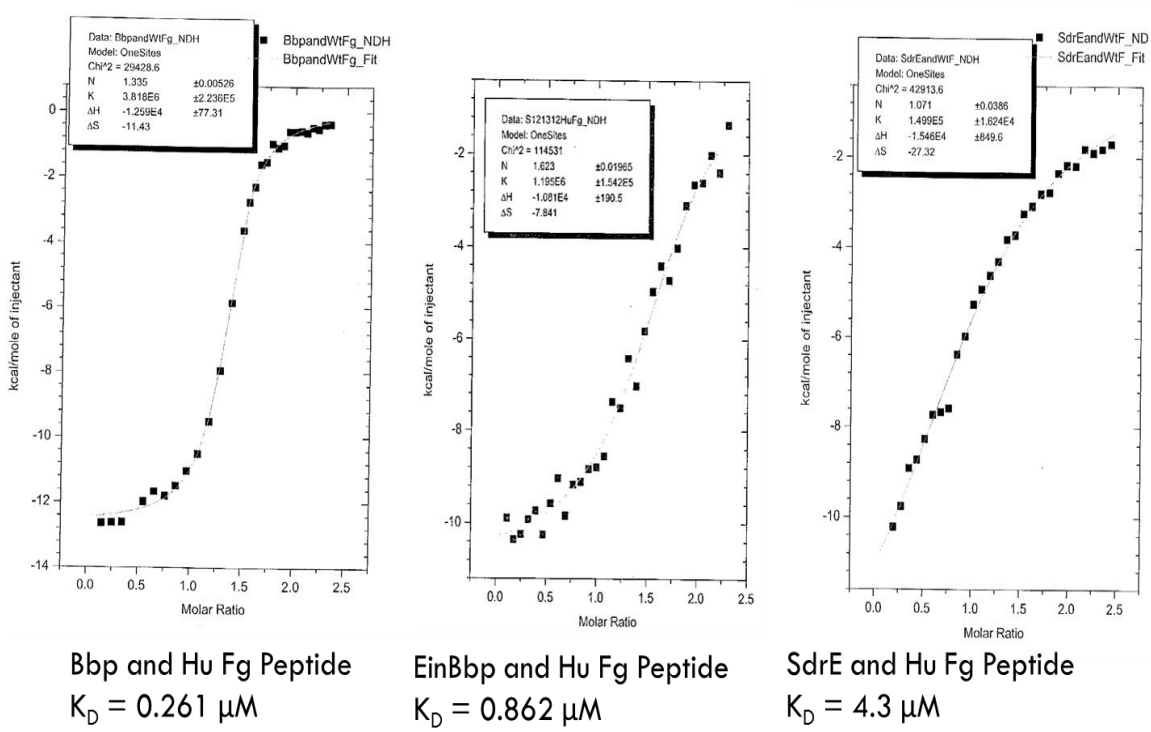


Figure 5-9. ITC measurement of SdrE2-LockChimera binding human fibrinogen peptide

The ability of SdrE2-LockChimera to bind to chips coated with fibrinogen from multiple species was tested in SPR experiments by Xiaowen Liang (Figure 5-10). SPR data confirmed observations from ELISA-type assay and ITC regarding SdrE2-LockChimera binding to human fibrinogen with an affinity between that of SdrE1-N2N3 and SdrE2-N2N3. Additionally, the SPR data confirms the ability for SdrE2-LockChimera to bind fibrinogen from multiple species. The ability of SdrE2-LockChimera to interact with dog, cat, and pig

fibrinogen shows that the Lock region of these two allelic variants is predominantly, but not entirely, responsible for the difference in specificity that the N2N3 constructs of these variants display (Figure 5-11).

Structures of apo-SdrE1-N2N3 and SdrE1-N2N3-Fg peptide co-crystal

Recently, the crystal structures of apo-SdrE1-N2N3 and SdrE1-N2N3-Fg peptide co-crystal have been solved by Ganesh Vannakambadi (Figure 5-12). The structures are highly similar to that of apo-SdrE2-N2N3 and the SdrE2-N2N3-Fg Peptide, respectively. These data confirm that SdrE1-N2N3 binds the Fg peptide via the Dock, Lock and Latch model. Moreover, the ligand binds to SdrE1 through an anti-parallel β-sheet interaction in the ligand binding trench. Key differences were discovered in the interaction of the allelic variants with the Fg peptide; these differences putatively explain the difference in binding profile between the two allelic variants.

The SdrE2-N2N3-Fg Peptide co-crystal reveals that Fg Aα 561-572 represent the true binding site with residues 573-575 (Arg-Gly-Asp) being outside of the ligand binding trench. Furthermore, 6-7 hydrogen bond formations between the backbones of the peptide and MSCRAMM. In contrast, in the SdrE1-N2N3-Fg Peptide co-crystal, the true binding site only contains Fg Aα 561-570 interacting in the ligand binding trench.

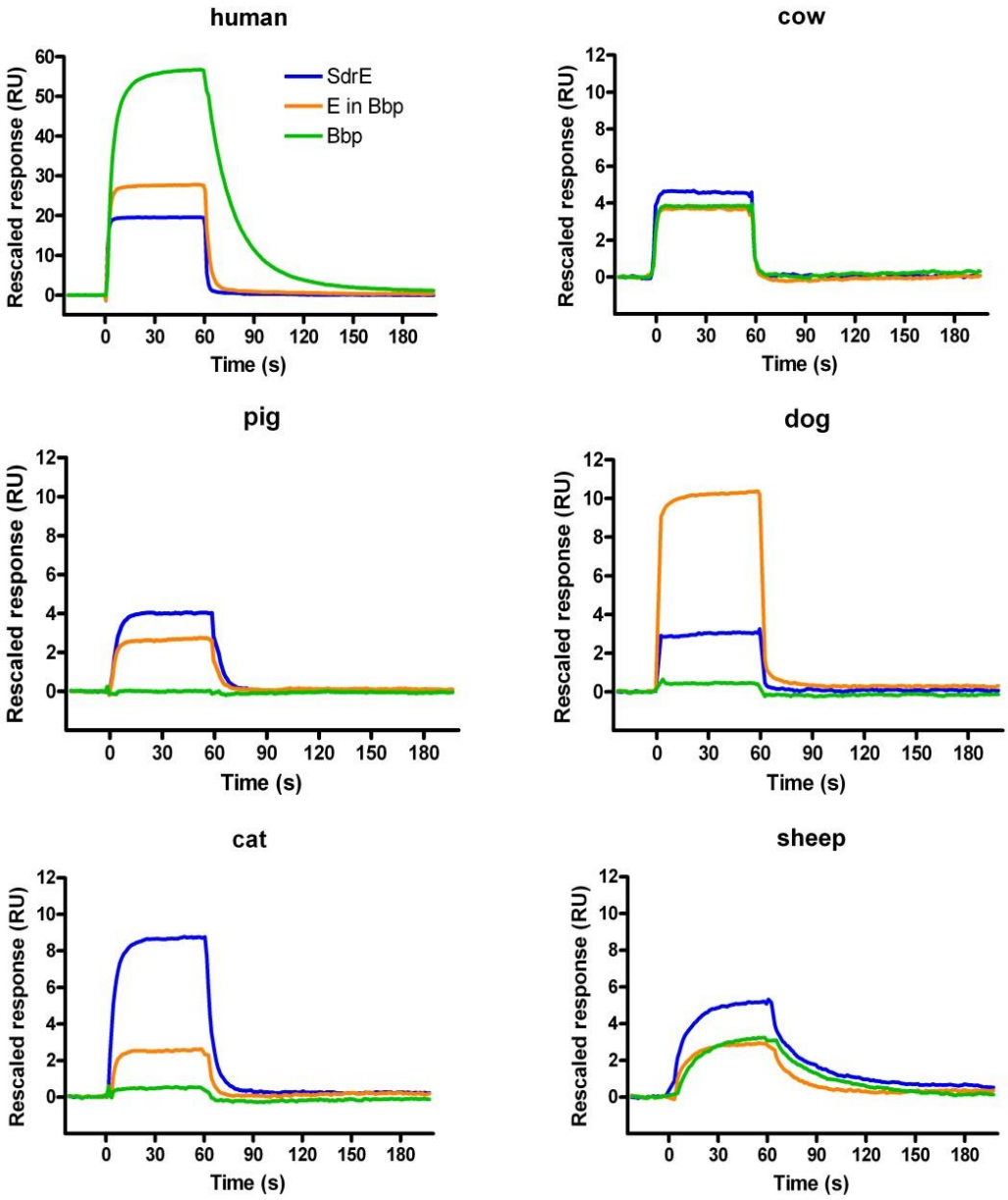


Figure 5-10. SPR measurement of SdrE2-LockChimera binding coated fibrinogen from various species.

SdrE2-LockChimera (orange) was tested in the same kinetic experiments as SdrE1-N2N3 (blue) and Bbp/SdrE2-N2N3 (green).

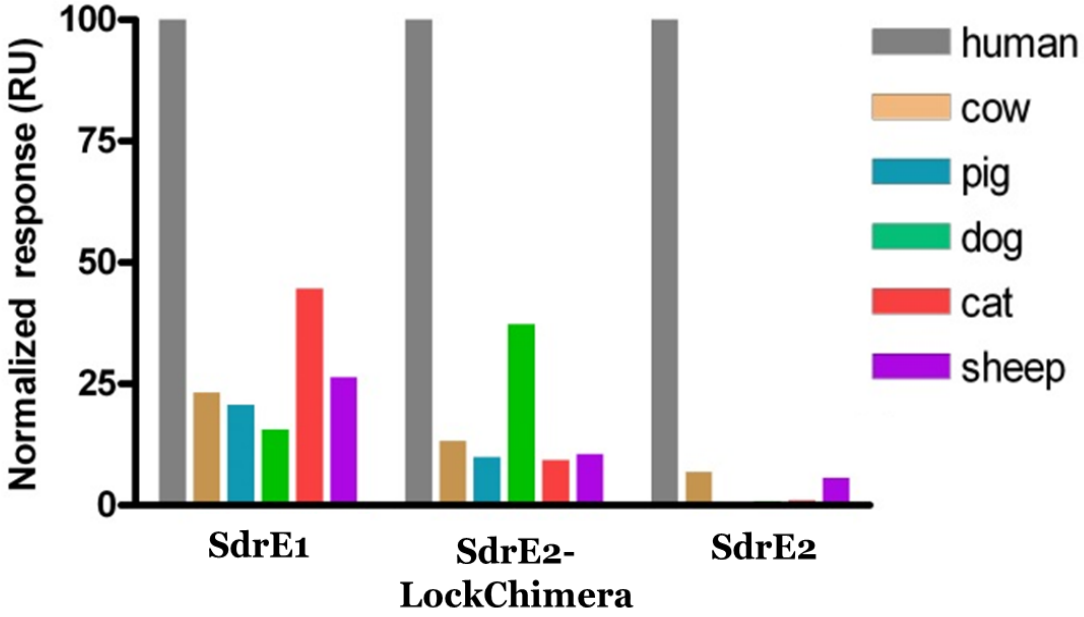


Figure 5-11. Summary of SPR data.

Furthermore, only 4-5 hydrogen bonds are formed between the peptide and MSCRAMM backbones. The reduced number of interacting residues are predicted to account for part of the reduction in affinity seen between SdrE1-N2N3 and human fibrinogen as compared to SdrE2-N2N3. This will be confirmed at a later date with experiments using a peptide that represents this shorter sequence.

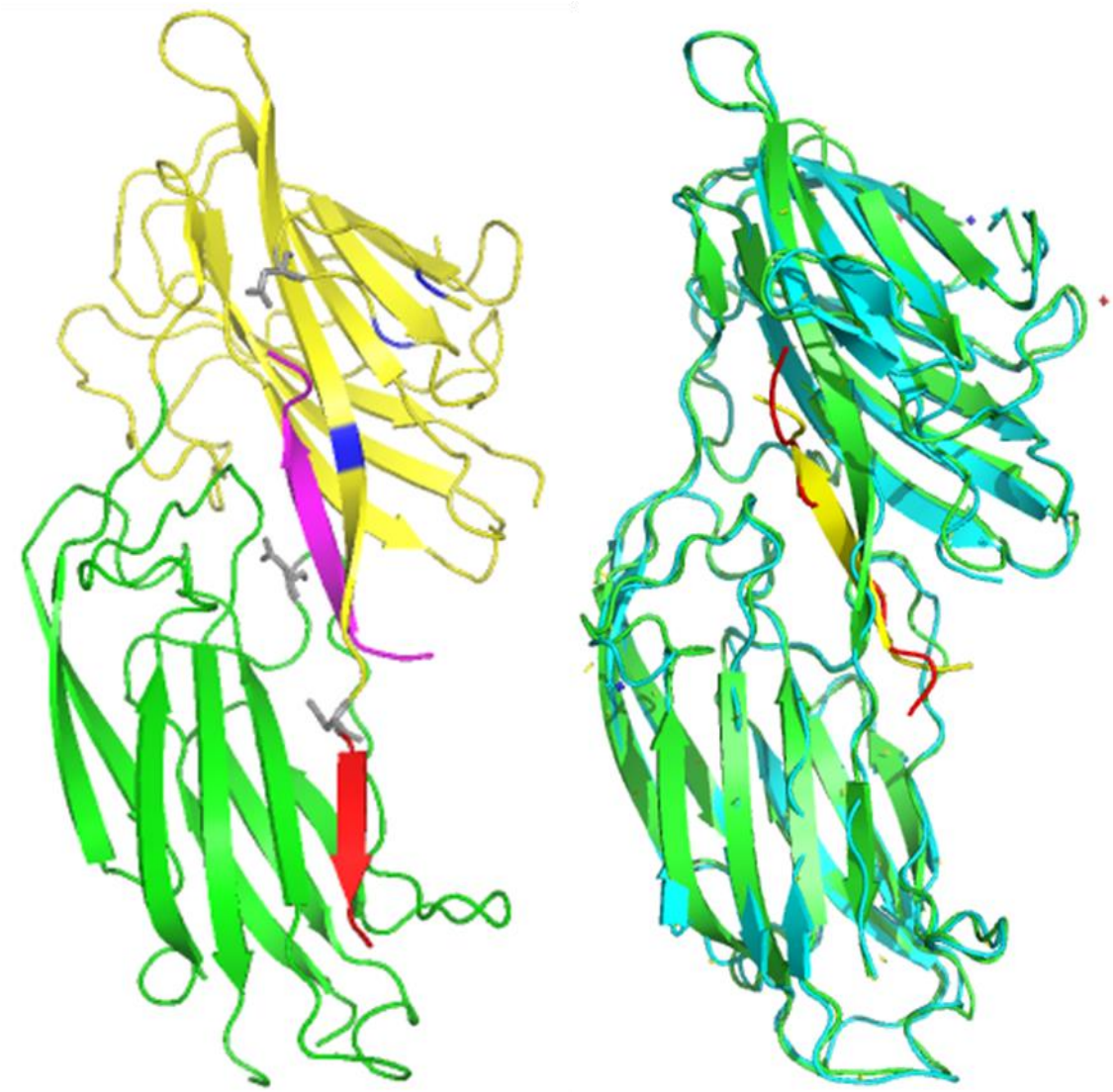


Figure 5-12. Structure of SdrE1-N2N3-Fg Co-crystal.
Panel 1 shows the structure of SdrE1-N2N3 bound to the human fibrinogen peptide. The N2 domain (green), N3 domain (yellow), fibrinogen peptide (purple), and Latch (red) domains are shown. Panel 2 shows an overlay of the fibrinogen-bound structures of SdrE1-N2N3 (green) and SdrE2-N2N3 (cyan).

Mutational analysis of SdrE1-N2N3

In order to confirm that the residues isolated for their putative role in the binding of SdrE1-N2N3 to fibrinogen play an important role, point mutations will be made in SdrE1-N2N3. SdrE1-D334A and SdrE1-D558A constructs will be made using site directed mutagenesis as described previously. After expression and purification, these mutants will be tested for binding to fibrinogen in ELISA-type assay. The hypothesis is that neither mutant will be able to bind to fibrinogen.

Progress towards a chimeric SdrE1 with the binding profile of SdrE2

Significant progress has been made towards understanding which residues of SdrE1 are needed in SdrE2 to change the binding profile to more closely resemble SdrE1. Namely, the replacement of the Lock region of SdrE2 with the SdrE1-Lock region results in an SdrE2-LockChimera that binds human fibrinogen with lower affinity and is able to show reduced binding to fibrinogen from multiple species. However, attempts at making a chimeric SdrE1-N2N3 that has the binding profile of SdrE2-N2N3 have been unsuccessful. The Lock region of SdrE1 was replaced with the SdrE2-Lock region, but the resulting SdrE1-LockChimera showed a reduced affinity for human fibrinogen instead of increased affinity and still bound to fibrinogen from multiple other species (Figure 5-13).

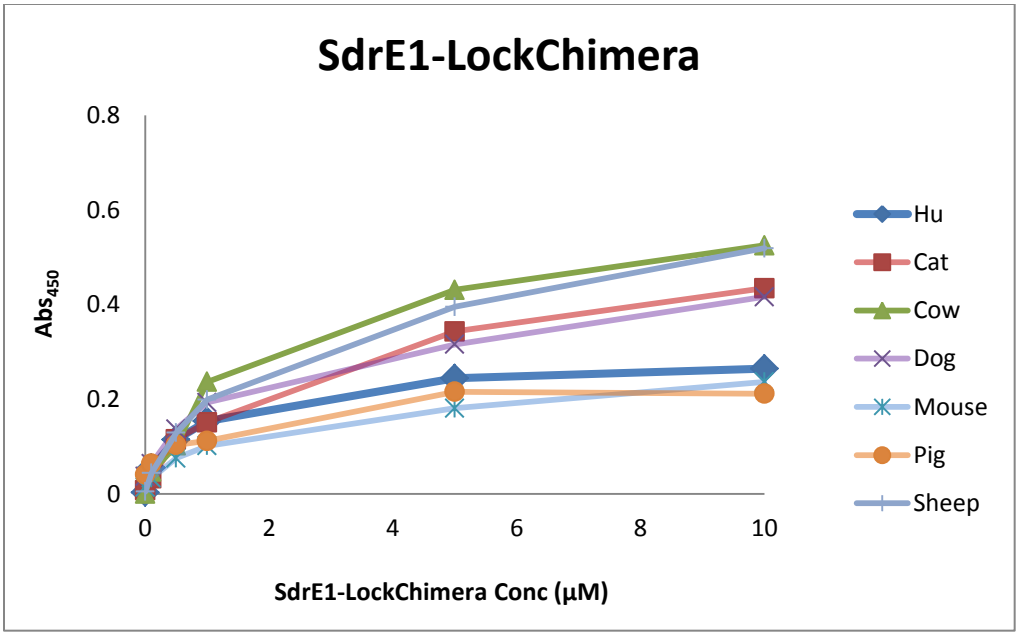


Figure 5-13. SdrE1-LockChimera animal fibrinogen binding profile.

Data generated from the solved structures of apo-SdrE1-N2N3 and SdrE1-N2N3-Fg Peptide co-crystal allows for a greater understanding of the structural basis for the differences in binding. Based on these analyses, there are three residues that could together be responsible for the differences in binding. In the binding of SdrE2-N2N3 to human fibrinogen, SdrE2-Y475 interacts with the peptide backbone at FgAα-S570. In SdrE1, this residue is an alanine, which prevents this hydrogen bond from forming. Instead, SdrE1-Q467 forms the requisite hydrogen bond, but the change in residue results in a shift of the rest of the ligand peptide within the ligand binding trench, preventing further contacts downstream of the S570 residue. By making both A475Y and Q467G point

mutations in the SdrE1-N2N3 construct, this combination mutant should be able to have a greater number of contact points than the wild type. Additionally, a F582T mutation will allow for additional hydrogen bond interaction with the Fg peptide. In sum, it is hypothesized that an SdrE1-Chimera will be made that shows a fibrinogen binding profile that closely resembles SdrE2-N2N3.

DISCUSSION

Previously, we showed that the ligand binding domains allelic variants SdrE1 and SdrE2 have dramatically different fibrinogen binding profiles despite being highly similar. There are multiple approaches that could have been used to find the underlying structural basis for the observed biochemical phenotypes. While these variants are 67% identical by amino acid sequence in the ligand binding domains, a library of chimeric point mutants would need to contain over 100 separate mutations from SdrE1 to SdrE2 and another 100 mutations from SdrE2 to SdrE1.

Furthermore, it was likely that we would need to combine multiple mutations in order to create a chimeric SdrE1 or SdrE2 that displayed the phenotype of the other variant. Instead, we have used the solved crystal structures of the ligand binding domains of SdrE1 and SdrE2, in complex with their target sequence in fibrinogen, to understand the structural basis of the functional differences in binding between SdrE1 and SdrE2. While difficult to

obtain, the work by Dr. Ganesh Vannakambadi to solve these structures have allowed for a much greater understanding of the mechanism by which these SdrE variants binds fibrinogen and identifying the putative residues that are critical for binding. The highly similar amino acid sequences of the allelic variants resulted in structures that had many similarities; however, upon closer inspection, key differences were observed.

Mutational analysis of the target site in fibrinogen has shown that the valine insertion at position 565 significantly reduces the binding capacity of SdrE2-N2N3. While the insertion shifts the remaining residues one position, our work shows that much of the reduction in binding is caused by the placement of the hydrophobic valine in the position where the hydrophilic threonine is located in the interaction with human fibrinogen. Additionally, a smaller reduction in binding occurs because of the placement of a glycine where an arginine normally occurs. These findings strongly suggest that if a mutation by the human host were to occur in this region, SdrE2 would more likely to be effected than SdrE1 due to the high specificity of SdrE2 for the wild type sequence.

Mutational analysis was used to confirm that residues predicted to play a role in binding by the solved crystal structure are critical for binding *in vitro*. These studies allowed for the eventual discovery of a chimeric SdrE2-N2N3 construct that displayed an intermediate fibrinogen binding profile when compared to the binding profiles of the two allelic variants. This construct was made replacing the Lock region of SdrE2 with the Lock region of SdrE1. The

SdrE2-LockChimera displayed a reduced affinity for human fibrinogen and a broad specificity for fibrinogen from non-human species in both ELISA-type assay and SPR experiments. While the SdrE1-LockChimera did not result in a similar intermediate phenotype, the recently solved SdrE2-N2N3-Fg Peptide structure has provided a handful of putative residues that could allow for further mutations to be made to create a functionally chimeric SdrE2.

It is important to note that a small number of changes in amino acid sequence result in a large change in phenotype. The fact that SdrE2-N2N3 displays a markedly higher affinity and specificity for human fibrinogen and the fact that there are less variations of *sdrE2* in the public database suggest that SdrE2 could be an example of host adaptation by *S. aureus*. While *S. aureus* is mainly a human pathogen, understanding how the bacteria can adapt to other hosts has been of significant interest in the field.

Specifically, mutations in SdrE2 resulted greater binding to human fibrinogen; this improvement in function would likely have come at the cost of binding to fibrinogen from most animals. There are more changes in amino acid sequence between SdrE1 and SdrE2 than are necessary for this change in phenotype, but this also speaks to the high frequency of mutation in this region. This high frequency of mutation could pose a health risk to humans in light of the minimal number of changes necessary for a change in phenotype. It is apparent that close monitoring of these regions of *Staphylococcus aureus* is

needed in order to monitor for the emergence of new, highly virulent allelic variants.

Whole genome sequencing is a powerful tool that potentially allows for a greater understanding of the evolutionary changes that are occurring within *S. aureus*. However, *in silico* analysis still needs to be paired with *in vitro* and *in vivo* research in order to fully maximize the potential of these data. The gain or loss of genes and the introduction of stop codons can be detected through sequence analysis; however, other variations may be functionally relevant yet not readily apparent.

Allelic variants SdrE1 and SdrE2 provide an example of the latter. With the information gained about residues critical for binding in these studies, variations in sequence of *sdrE* can be modeled to project which sequence changes could result in changes in binding.

CHAPTER VI

BBP/SDRE2 AND BONE SIALOPROTEIN BINDING

INTRODUCTION

Osteomyelitis and *Staphylococcus aureus*

Osteomyelitis is an infection of the bone that can be caused by bacteria or fungi and occurs in approximately 2 out of every 10,000 people. It is generally classified by duration, pathogenesis, site, extent and/or type of patient. Mixed infections occur frequently enough that osteomyelitis infections are not classified by pathogen.[8]

There are three major routes of infection for osteomyelitis: primary direct infection via trauma or invasive medical procedure, secondary infection via the bloodstream, or secondary infection via spread from local tissue infection. The most common route of infection depends on the age and comorbidities of the patient. In children, hematogenous osteomyelitis is the most common route of infection and the long bones are the most common site of infection. It is hypothesized that the growing bones of children are more porous to bacteria from the vasculature than mature bones in adults.[65] Hematogenous osteomyelitis can occur in adults, but more often results in an infection of the vertebra. In adults, primary direct infection and secondary infection via spread from local soft tissue infection are more common than secondary hematogenous

osteomyelitis. Primary osteomyelitis is a concern in trauma and surgical procedures, especially when implants are placed into the bone tissue. Secondary osteomyelitis from local tissue is of particular concern in chronic diabetic patients due to the impaired neurovasculature in the distal extremities resulting in poor blood flow and reduced sensation. Chronic granulomatous disease and sickle cell anemia are also risk factors for osteomyelitis. [9,65]

While mixed infections are common in osteomyelitis, *Staphylococcus aureus* is the leading cause of osteomyelitis across all age groups and comorbidities. Approximately 50-70% of osteomyelitis infections contain *S. aureus*. The rise of methicillin-resistant *S. aureus* has been cited as the biggest epidemiologic challenge in osteomyelitis and osteoarticular infections.[8]

Bbp

In 1989, it was published by Rydén *et al* that *S. aureus* isolates from osteomyelitis patients are able to bind bone sialoprotein (BSP).[52,66] Further work by these investigators isolated an MSCRAMM as the surface factor responsible for this binding. This MSCRAMM was named Bone sialoprotein–Binding Protein, Bbp.[53] The interaction between Bbp and BSP was further isolated to the nonapeptide sequence, LKRFPVQGG, that occurs in the N-terminal half of BSP.[46] Additional evidence for the role of Bbp in staphylococcal osteomyelitis is seen in molecular epidemiological studies that show *bbp* to be overrepresented

in osteomyelitis strains compared to staphylococcal strains from all pathologies.[54] An interesting study has shown a proof of concept for a technique that allows for discrimination between SSTI and osteomyelitis in diabetic foot infections. The often difficult task is accomplished via serological assay that probes for patient antibodies to Bbp. [67]

Bone sialoprotein

Bone sialoprotein (BSP) is a 301 amino acid, highly flexible, extracellular matrix protein that is expressed by osteoblasts in bone tissue. Bone tissue is made up of two integral components – collagen, which provides elasticity and flexibility – and hydroxyapatite mineralization, which provide strength and rigidity. While collagen is by far the predominant protein in bone tissue, bone sialoprotein is one of the major non-collagen proteins.[68] Bone sialoprotein serves two major roles in bone tissue. First, it acts as a site of nucleation of hydroxyapatite crystal formation. Second, it allows for tissue specific cells to attach to the matrix.[69]

BSP has been extensively studied and reported on in the literature. BSP is divided into 3 sub-regions: a basic N-terminal region, a central domain, and an acidic C-terminal region. Within the second half of the central domain, there are a number of serine residues that are phosphorylated after BSP is secreted by osteoblasts into the bone matrix. The high density of charged phosphate

molecules in a concentrated region has been shown to serve as a nucleation site of hydroxyapatite crystal formation.[68] While other phosphoproteins function in hydroxyapatite nucleation, BSP is the most potent known nucleator of hydroxyapatite.[69] It has recently been discovered that phosphorylation of Ser136 is critical for the nucleation event.[70] The degree of mineralization of bone tissue can be regulated by the host via expression of extracellular kinases or phosphatases.

Given that hydroxyapatite has no binding sites for the cells of the bone tissue, the ability to bind bone sialoprotein allows these bone cells to bind regions of the tissue that are highly mineralized. This interaction occurs through a C-terminal RGD motif, a classical integrin-binding motif, as well as sulfated tyrosine residues. In addition to these structure-function motifs, there is large number of serine, threonine and tyrosine residues, which allow for extensive O-linked glycosylation, and asparagine residues, which allow for N-linked glycosylation.

Bone sialoprotein plays an important role in the generation of new bone tissue. In light of this, it is not surprising that the *Bsp* gene has been found to be upregulated in multiple different mouse models of bone fracture, ranging from non-displaced fractures to compound fractures. The expression of *Bsp* is upregulated 3-7 days after injury; this which could play a role as an important virulence mechanism for osteomyelitis-causing bacteria.[71] In adults, *Staphylococcus aureus* causes osteomyelitis in trauma cases where there is a

direct infection of the bone tissue. Given that *S. aureus* is an opportunistic pathogen, the ability to target a host protein that is highly upregulated in trauma would be of great benefit to the bacteria.

Conflict in the literature

Since the publication of SdrE2 binding BSP, the Höök lab has made multiple attempts at replicating these findings, but has been unsuccessful. The recent publication showing that SdrE2 binds human fibrinogen creates conflicting mechanisms of virulence function of SdrE2. The original publications found a specific interaction of SdrE2 with bone sialoprotein and no interaction with fibrinogen, amongst other ECM proteins.[53] Our more recent findings show that SdrE2 does bind to fibrinogen, but have been previously unsuccessful at showing an interaction with BSP. [47]

RESULTS

Peptide analysis of putative target site of SdrE2 on bone sialoprotein

Given that previous attempts to show binding between SdrE2-N2N3 and full length bone sialoprotein were unsuccessful, we attempted to determine a structural-based rationale for binding or lack of binding to BSP using peptides representing the published binding sites of human fibrinogen, bone sialoprotein

or chimeras of the two. The nonapeptide sequence that Bbp/SdrE2 was reported to target in BSP (LKRFPVQGG) bore some similarity to the targeted sequence in fibrinogen, especially when comparing the first four residues.[46] It was determined that the first four residues of the BSP sequence were the most important for binding. This is interesting in light of the similarity between the first four residues of the human fibrinogen target sequence and the bone sialoprotein target sequence.

Furthermore, when this sequence was queried in the public database, it was discovered that the amino acid sequence in this region of BSP is actually is LKRFPVQGSSDSS. This contains a single amino acid change from the reported sequence. Additionally, we observed that the four residues that occur after the reported nonapeptide sequence result in a peptide sequence that is more similar to the fibrinogen sequence (Figure 6-1, Top).

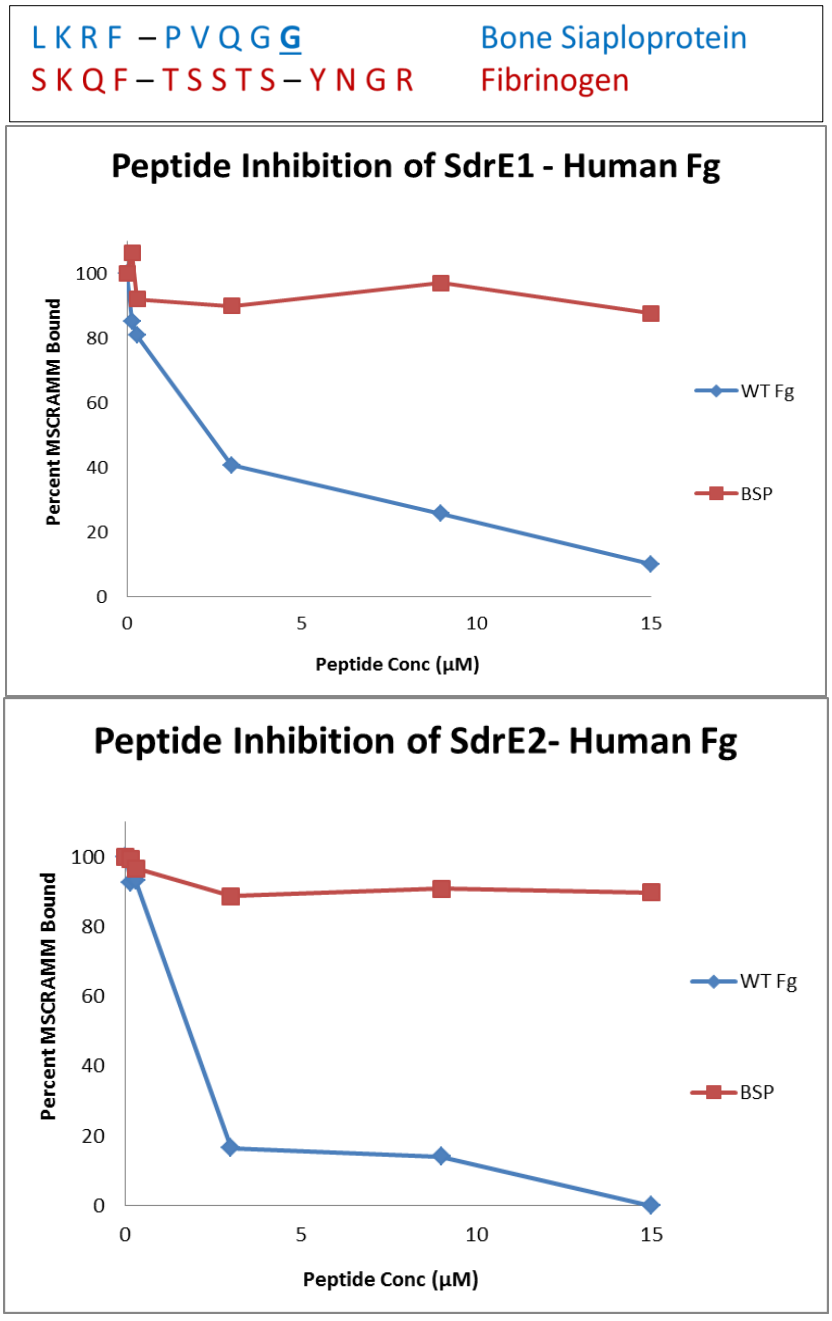
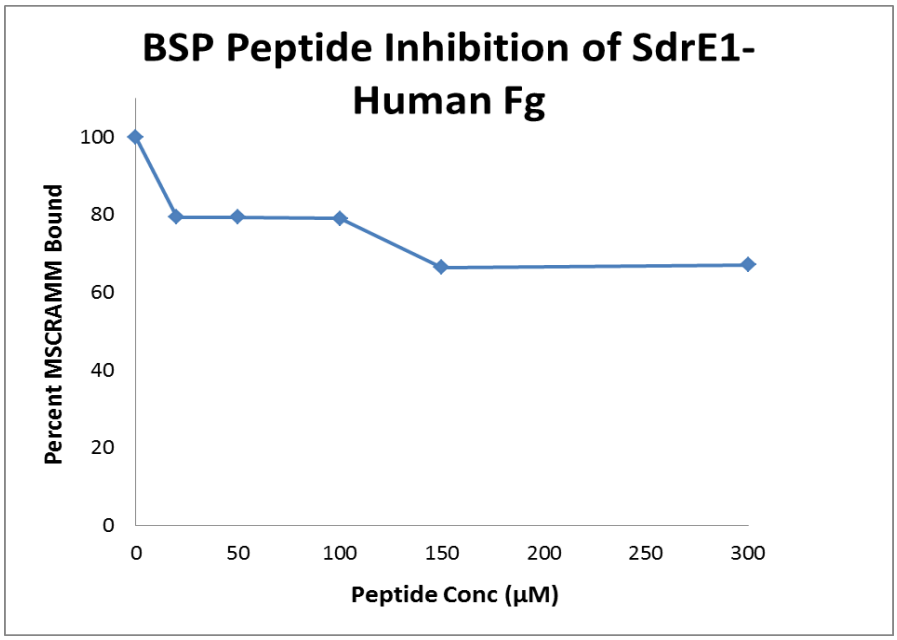


Figure 6-1. Attempts at BSP peptide inhibition of SdrE-Fibrinogen.
Top panel shows alignment of BSP and Fg target sites. Bottom panel shows results of Peptide Inhibition ELISA-type assays wherein the published peptide cannot inhibit binding of SdrE to fibrinogen.

Peptide Inhibition ELISA-type assays were performed in which 0.3 μM SdrE1 or SdrE2 was incubated with increasing amounts of the designated peptide before this mixture was added to fibrinogen-coated wells. If the protein binds the peptide, there is a resultant reduction in signal due to an inability for the peptide-bound MSCRAMM to subsequently bind to the fibrinogen-coated wells.

While the Fg peptide was able to inhibit binding of SdrE1-N2N3 or SdrE2-N2N3 to fibrinogen-coated wells, the BSP peptide was unable to do the same in the initial assay measuring up to a 1:50 ratio of MSCRAMM:Peptide (15 μM) (Figure 6-2). The experiment was repeated with a concentrations reaching an upper limit of 1:1000 (300 μM); SdrE2-N2N3 showed no affinity for this peptide. SdrE1-N2N3 showed some interaction, but the concentrations required to achieve greater than 20% inhibition suggests that this interaction is of a very low affinity.



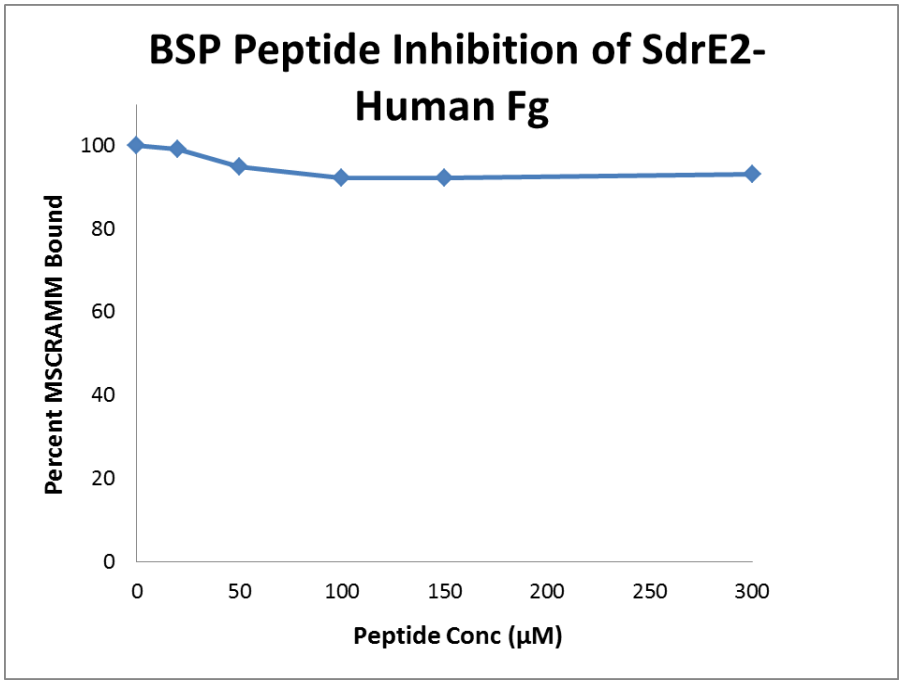
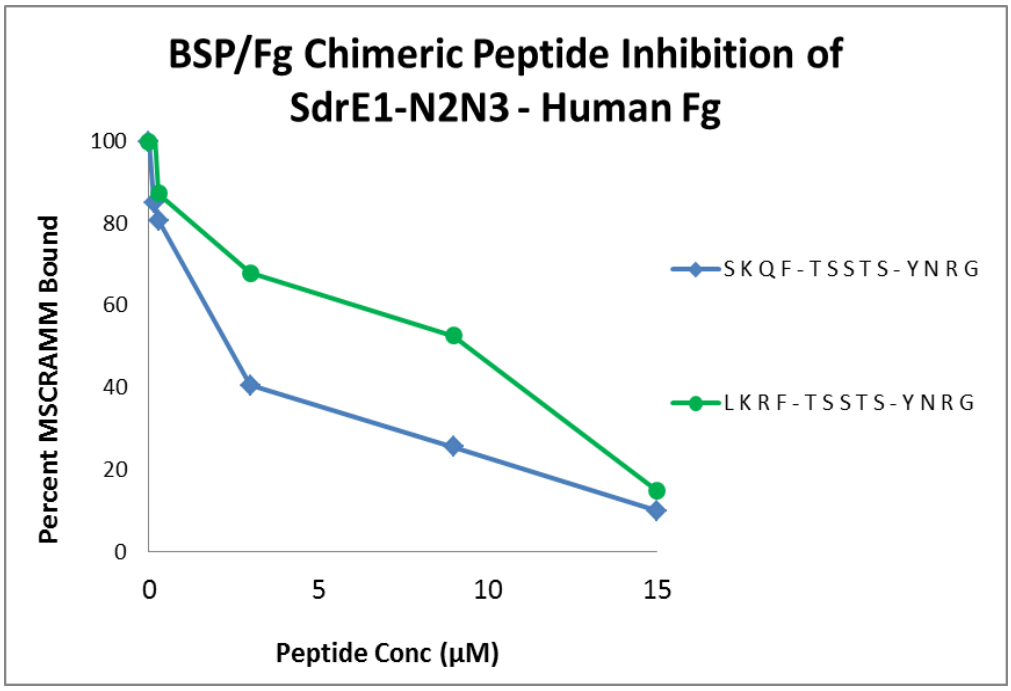


Figure 6-2. Attempts at BSP peptide inhibition of SdrE-Fibrinogen.
Results of Peptide Inhibition ELISA-type assays wherein the published peptide can not inhibit binding of SdrE to fibrinogen.



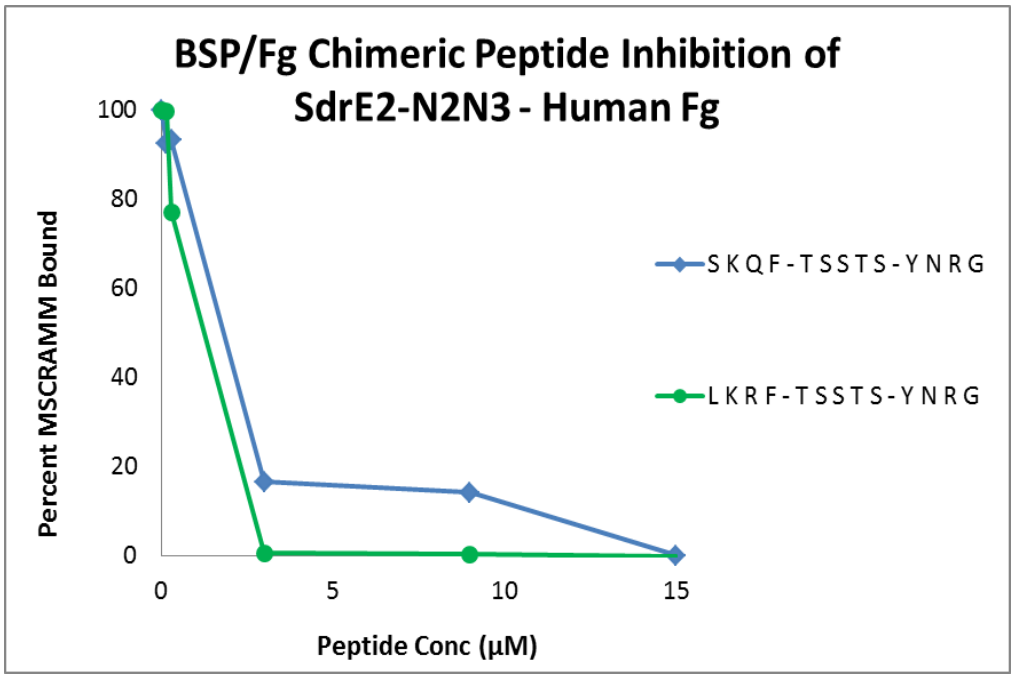


Figure 6-3. Peptide inhibition of SdrE-Fibrinogen with chimeric peptide.

Interestingly, a chimeric peptide made of the first four residues of BSP and the last 9 residues of Fg is not only bound by SdrE2-N2N3 (Figure 6-3), but is bound with an apparent affinity that is even greater than the apparent affinity for the wild type Fg peptide. This peptide was made based on the finding that the first four BSP residues were critical to binding and also our observation that these residues are similar to the first four residues of the Fg peptide.

The resulting peptide has only two changes from the Fg peptide. These findings were confirmed using ITC with the chimeric peptide and SdrE2-N2N3. In these experiments, there was a dramatic increase in affinity when compared to the wild type fibrinogen peptide (Figure 6-4). We were surprised to note that SdrE2-N2N3 displayed a 10-fold greater affinity for this chimeric peptide than it did for the wild type human fibrinogen peptide despite the minimal changes. This raises the possibility that the binding interactions gained in the first four residues could overcome other problematic interactions in the C-terminal half of the binding sequence.

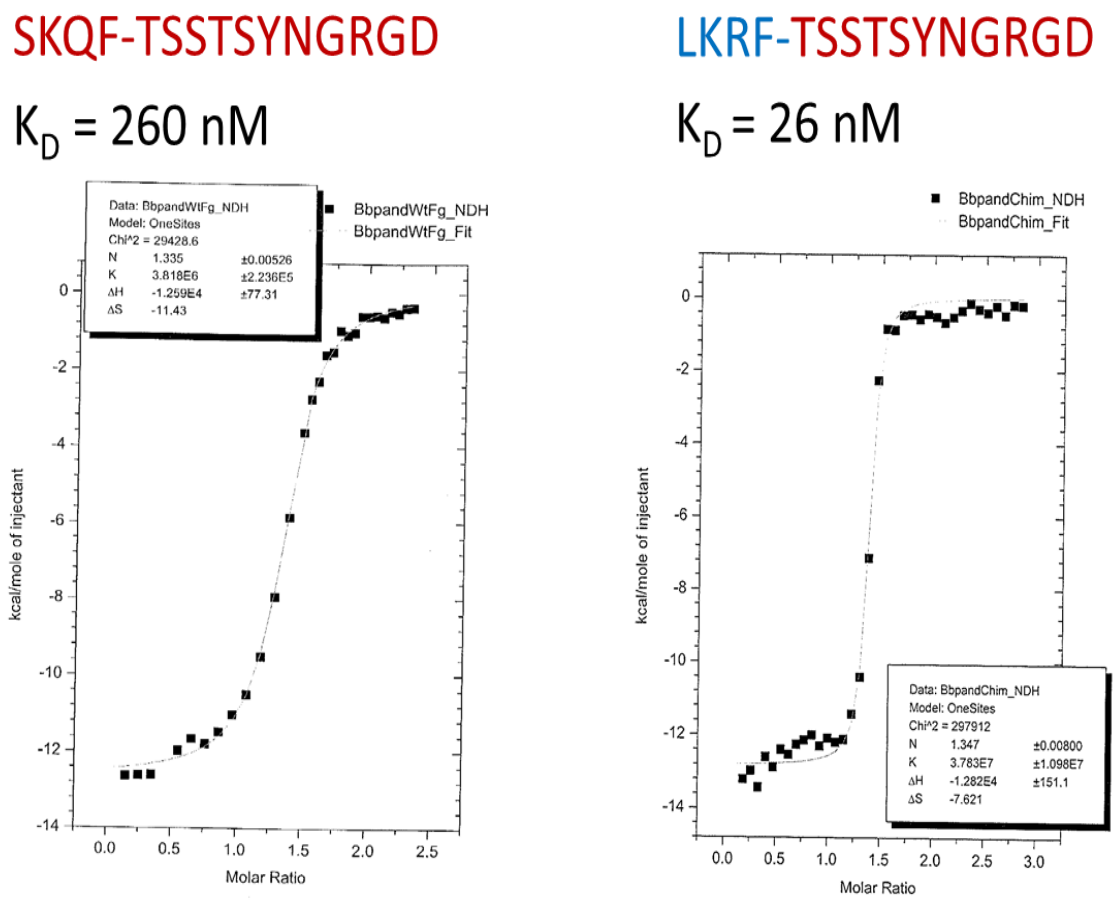


Figure 6-4. ITC measurement of SdrE2-N2N3 binding WT and Chimeric Peptide.

Panel 1 shows the binding curve of the wild type human fibrinogen peptide injected into SdrE2-N2N3. Panel 2 shows the injection of a BSP-Fg chimeric peptide injected into SdrE2-N2N3.

Previously, mutational analysis of the Fg peptide showed that certain residues are of critical importance in binding to SdrE2. In this case, we had evidence that it was possible to improve the binding interaction between SdrE and its ligand, but we had yet to understand the structural basis. To accomplish this, chimeric peptides were ordered that contained the changes as noted in Table 6-1. These peptides were then tested in Peptide Inhibition ELISA-type assays as described previously (Figure 6-5).

In addition to lack of binding to the BSP wild type sequence, SdrE2-N2N3 showed no binding to a P->T point mutant or P->T, S->R double mutant. These mutations were made because they were shown to affect binding when altered in the Fg sequence. Despite the importance of these residues in Fg binding, they were not enough to create an interaction with SdrE2-N2N3. While the S->L point mutant resulted in a lesser apparent affinity than the wild type Fg peptide, the Q->F mutation resulted in a greater apparent affinity.

Table 6-1. Bone Sialoprotein - Fibrinogen Chimeric Peptides

L K R F – P V Q G **G**	BSP – BSP –
S K Q F – T S S T S – Y N G R	Fg – Fg – Fg
L K R F – T S S T S – Y N G R	BSP – Fg – Fg
L K R F – P V Q G **S** – S D S S	BSP – BSP – BSP
L K R F – T V Q G **S** – S D S S	BSP – BSP/Fg – BSP
L K R F – T V Q G **S** – S D S R	BSP – BSP/Fg –BSP/Fg
L K Q F – T S S T S – Y N G R	BSP/Fg – BSP – BSP
S K R F – T S S T S – Y N G R	BSP/Fg – BSP – BSP

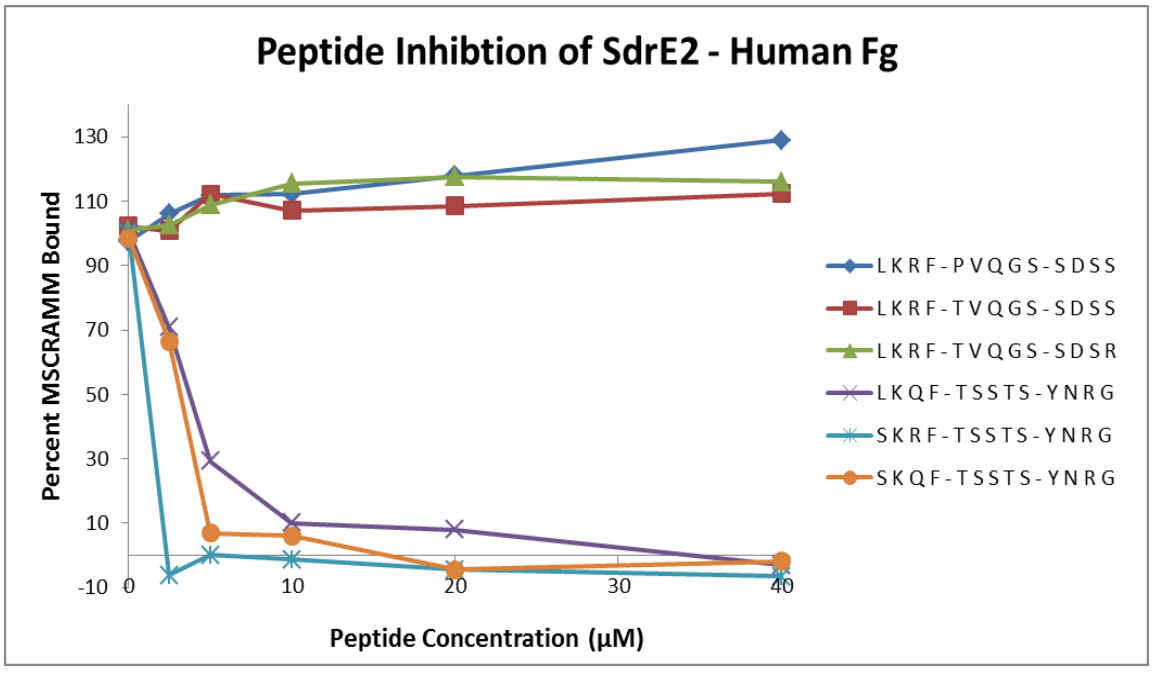


Figure 6-5. Chimeric Peptide Inhibition of SdrE2-N2N3 – Human Fibrinogen

Interestingly, this data corroborates previous experiments from the mutational analysis of SdrE2-N2N3 which showed that the D588 residue, located in the Lock-Latch region, is critical for SdrE binding. The SdrE2-D588A mutant showed no binding to fibrinogen. The solved SdrE2-N2N3-Fg Peptide structure provides a subatomic view of the ligand binding trench and the interactions with the peptide side chains and backbone. SdrE-D588 interacts with FgAα-Q563. By creating the FgAα-Q563R mutation, the negatively charged SdrE-D588 is putatively able to form a stronger bond with the positively charged arginine as opposed to the hydrogen bond it formed with the glutamine (Fig 6-5).

Lock variation in SdrE2-Rydén

In 2000, the Rydén lab published the sequence of the *sdrE2* that they had been using in their experiments. This sequence matches the sequence of the *sdrE2* used in the Höök lab, except for one 14 amino acid region (Figure 6-6). At the time, this region was not known to be important for binding other than being known to be located at the end of the A domain. However, this 14 amino acid mutation occurs in the Lock-Latch domain that our structure-function data now shows to be critically important for the binding of SdrE.

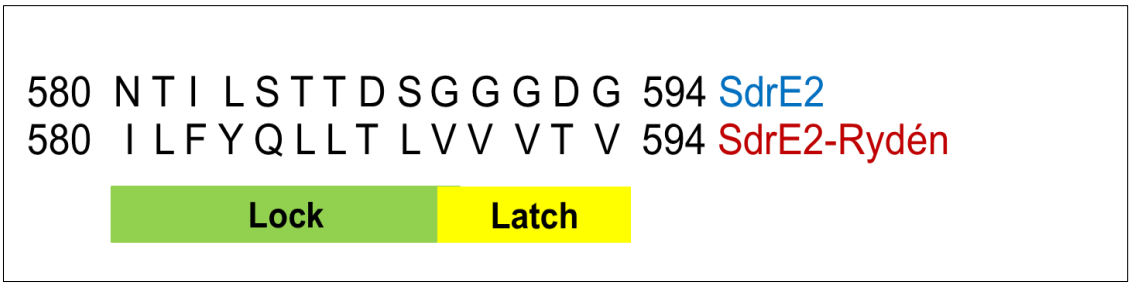


Figure 6-6. Lock-Latch Region Alignment of SdrE2

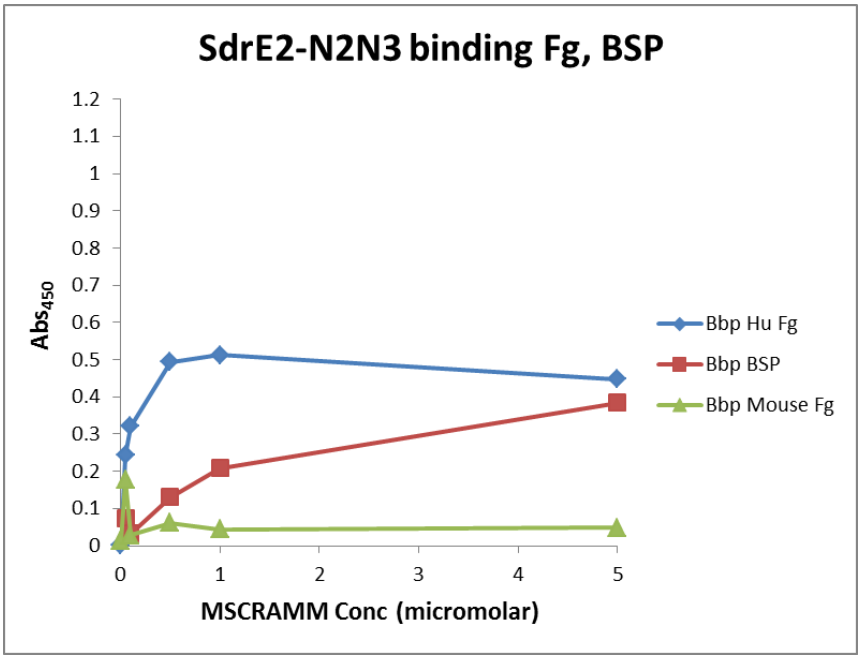
We have not observed this mutation in other SdrE sequences from publicly available databases or collaborators who work in the field of *S. aureus* genomics. This could mean that the variation observed by the Rydén lab is a cloning error or a sequencing error. However, this sequence comes from an osteomyelitis strain taken from a patient. It is possible that this isolate acquired a mutation that allowed for additional or altered functionality.

We hypothesized that, since the sequence of the SdrE2 that they use was derived from a staphylococcal osteomyelitis strain, they had discovered a variant of SdrE2 that does bind to BSP while the more commonly found SdrE2 that is used in the Höök Lab does not. To test this hypothesis, an SdrE2-Rydén-N2N3 construct was made by using a modified Overlap Extension PCR technique (see Chapter 2 - Methods).

SdrE2-Rydén-N2N3 and SdrE2-N2N3 were purified according to protocols previously described and then subsequently tested in ELISA-type

assays as previously described. In these assays, we were able to acquire bone sialoprotein that was purified from a CHO cell line from R&D Systems.

Surprisingly, wild type SdrE2-N2N3 is shown to bind to bone sialoprotein in the ELISA-type assays (Figure 6-7). This had yet to be demonstrated by the Höök lab; however, it is possible that this is able to be seen now because the BSP used here is derived from a human cell line which ostensibly results in a product that contains the many and varied types of glycosylation that bone sialoprotein has been shown to have *in vivo*. It is possible that these post-translational modifications play a direct role in binding of SdrE2 to BSP or result in an altered folded state of BSP that indirectly affects SdrE2 binding. Glycosylation has been shown to affect folding with many other proteins due to the ability of the sugar moieties to interact with the surrounding aqueous environment. Bone sialoprotein is a heavily glycosylated protein and these glycosylations are critically important to the function of the protein.



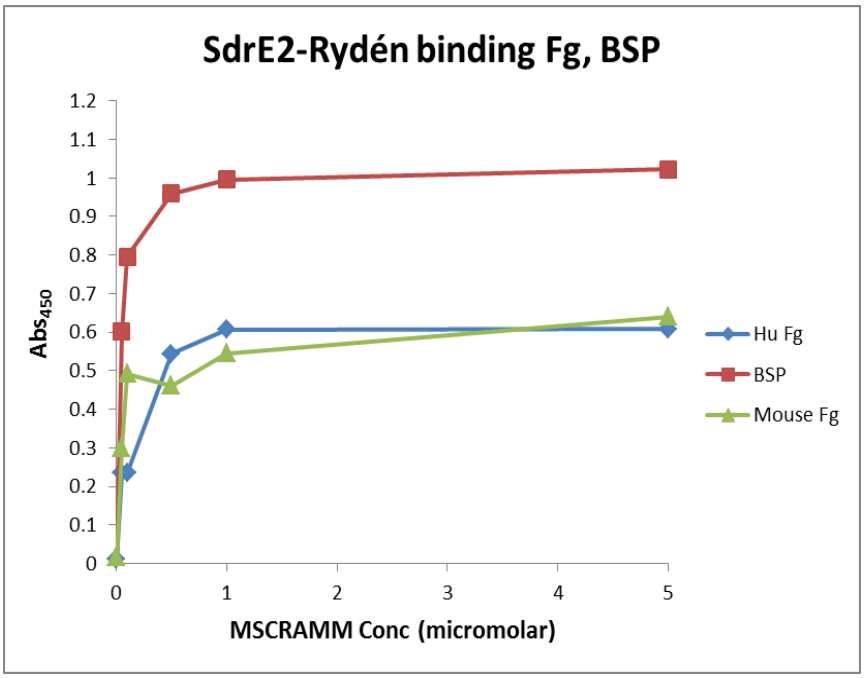
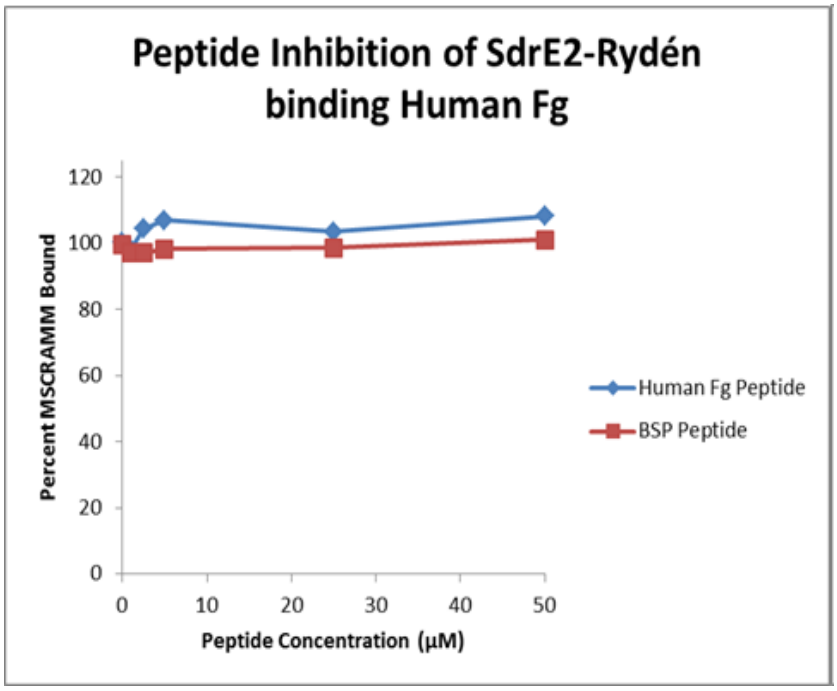


Figure 6-7. Bone Sialoprotein binding.
In panel 1, Wild Type SdrE2-N2N3 shows strong binding to Fg and some interaction with BSP. In panel 2, the variant SdrE2-N2N3 shows binding to everything.

Secondly, it appears as though SdrE2-Rydén binds not only to human fibrinogen, but BSP with a very high affinity. However, SdrE2-Rydén also appears to interact quite strongly with Mouse Fg, which was meant to serve as a negative control for this experiment. This finding raised concerns that the changes in SdrE2-Rydén resulted in a large degree of non-specific binding. In an attempt to address this, Peptide Inhibition ELISA-type assays were performed in order to see if binding by SdrE2-Rydén could be competitively inhibited, which would suggest a specific interaction (Figure 6-8).

Neither form of binding was inhibited, which raises the concern that the original observation in ELISA-type assay is could be attributed to non-specific binding. Additional concern comes from the knowledge that differences in sequence that have been introduced result in a much more hydrophobic set of residues in the Lock domain. This domain is supposed to be a flexible, hydrophilic domain that interacts with the aqueous environment before inserting into the N2 domain. The change to the more hydrophobic sequence of SdrE2-Rydén could have unintended consequences on the construct, potentially causing increased non-specific binding.



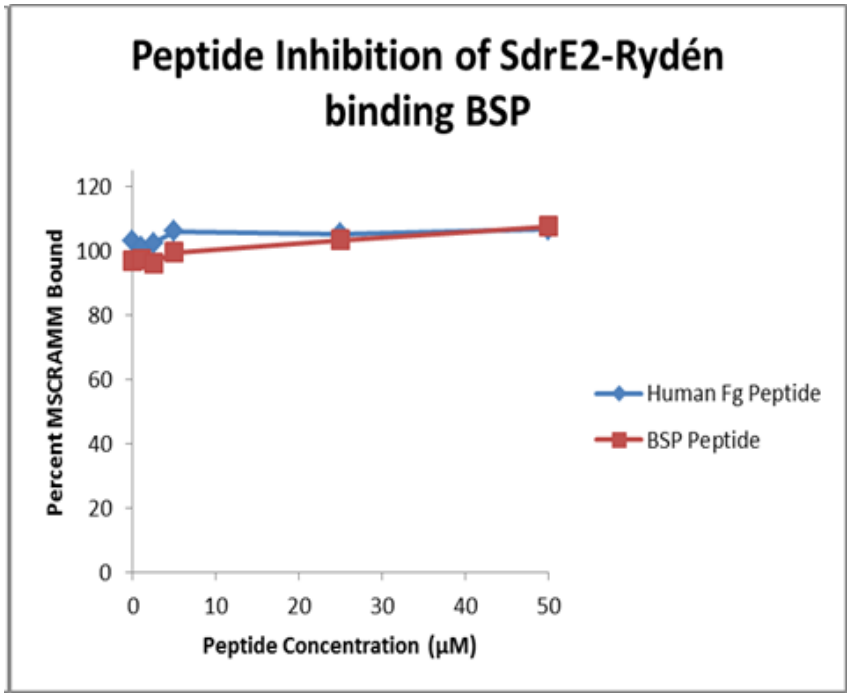


Figure 6-8. Attempts at Peptide Inhibition of SdrE2-Rydén.
Panel 1 shows lack of peptide inhibition of SdrE2-Rydén binding to fibrinogen while panel 2 shows lack of peptide inhibition of SdrE2-Rydén binding to bone sialoprotein.

SdrE2-Rydén forms putative dimers and multimers which display nonspecific binding

In subsequent purifications of the SdrE2-Rydén-N2N3 construct, we observed that peak fractions contained bands of a much higher molecular weight in 8% SDS-PAGE gels stained with Coomassie Blue (Figure 6-9). The predicted size of SdrE2-N2N3 is 34 kDa, which has been confirmed by mass spectrometry. In SDS-PAGE, SdrE2-N2N3 migrates between the 56 kDA and 45 kDA markers. SdrE2-Rydén-N2N3, being a mutant that has exchanged residues in the lock and latch region, has the same number of amino acids as well as the same predicted size. However, there was an observed band at approximately 80 kDa and another observed band that barely enters the gel, suggesting a very large size.

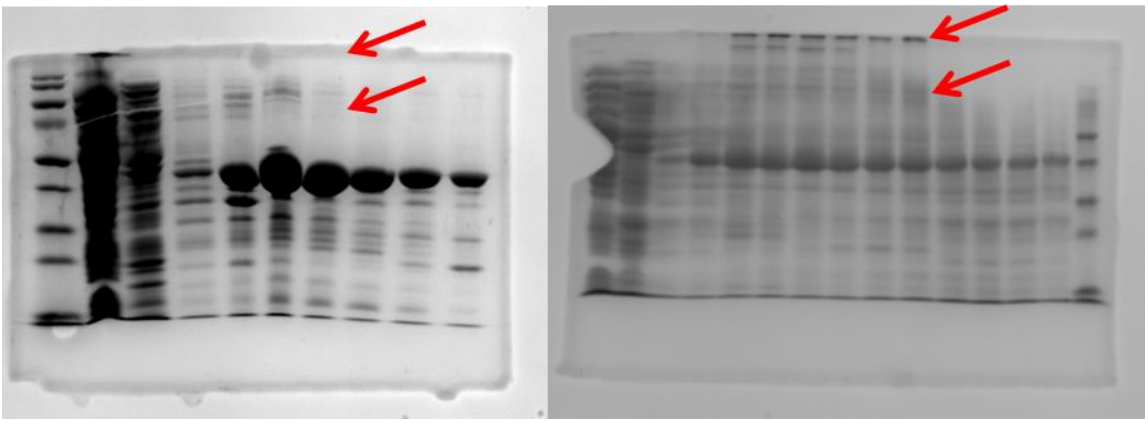


Figure 6-9. SDS-PAGE/Coomassie.
Panel 1 is a purification gel of SdrE2-N2N3. Panel 2 is a purification gel of SdrE-Rydén with the red arrows designating putative multimers.

There are no cysteine residues in the SdrE2-Rydén-N2N3 amino acid sequence which excludes the possibility that any putative dimerization or multimerization is occurring through disulfide bond formation. This was confirmed through running reducing and non-reducing gels (data not shown), which showed no difference in banding pattern. To further investigate the size and nature of these putative multimers, purified SdrE2-Rydén-N2N3 samples were purified through size exclusion chromatography using a 120 mL Sephacryl S-200 gel filtration column.

The SEC data confirmed that in addition to a monomer of approximately 34 kDa, SdrE2-Rydén-N2N3 also dimerizes and multimerizes (Figure 6-10). The largest peak occurs in the void volume of the column, consistent with a multimer of an indeterminate size. The putative dimer peak occurs at 75 kDa. Running the samples of a reducing SDS-PAGE gel showed the presence of a band that migrated at the size of the monomer, but this is likely an effect of the gel conditions. Importantly, the presence of a high molecular weight species representing the putative multimer and an 80 kDa molecular weight species representing the putative dimer are very faint in the fractions from the monomer peak. Each of the three peaks were collected and kept separately at 4˚C.

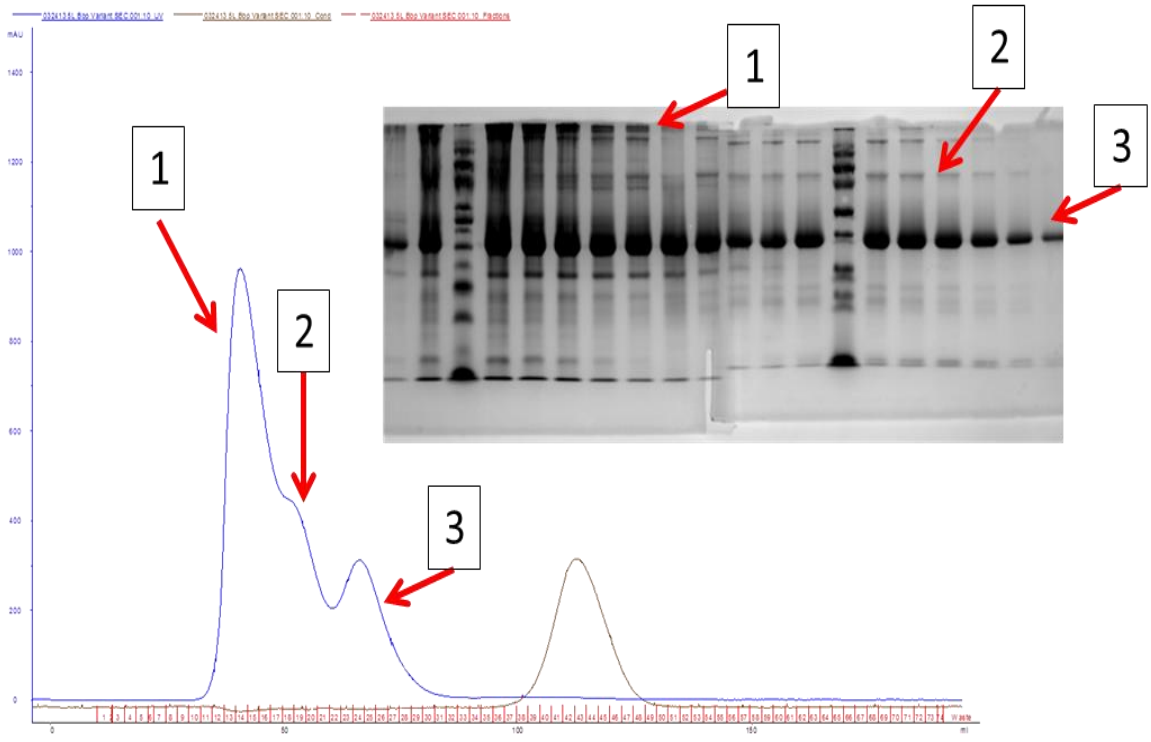


Figure 6-10. Size Exclusion Chromatography of SdrE2-Rydén.
The three peaks in SEC correlate to 3 bands on SDS-PAGE. 1: SdrE2-Rydén Multimer, 2: SdrE2-Rydén Dimer, 3: SdrE2-Rydén Monomer.

Given that this multimerization was occurring without disulfide bonds, it was hypothesized that the multimerization process may be a reversible process that could be driven towards the monomeric state via dilution. To test this, the multimer and dimer fractions were diluted 10-fold using 1x TBS buffer, and then slowly concentrated to 5 mL to run a second time on SEC on the 120 mL Sephacryl S-200 column.

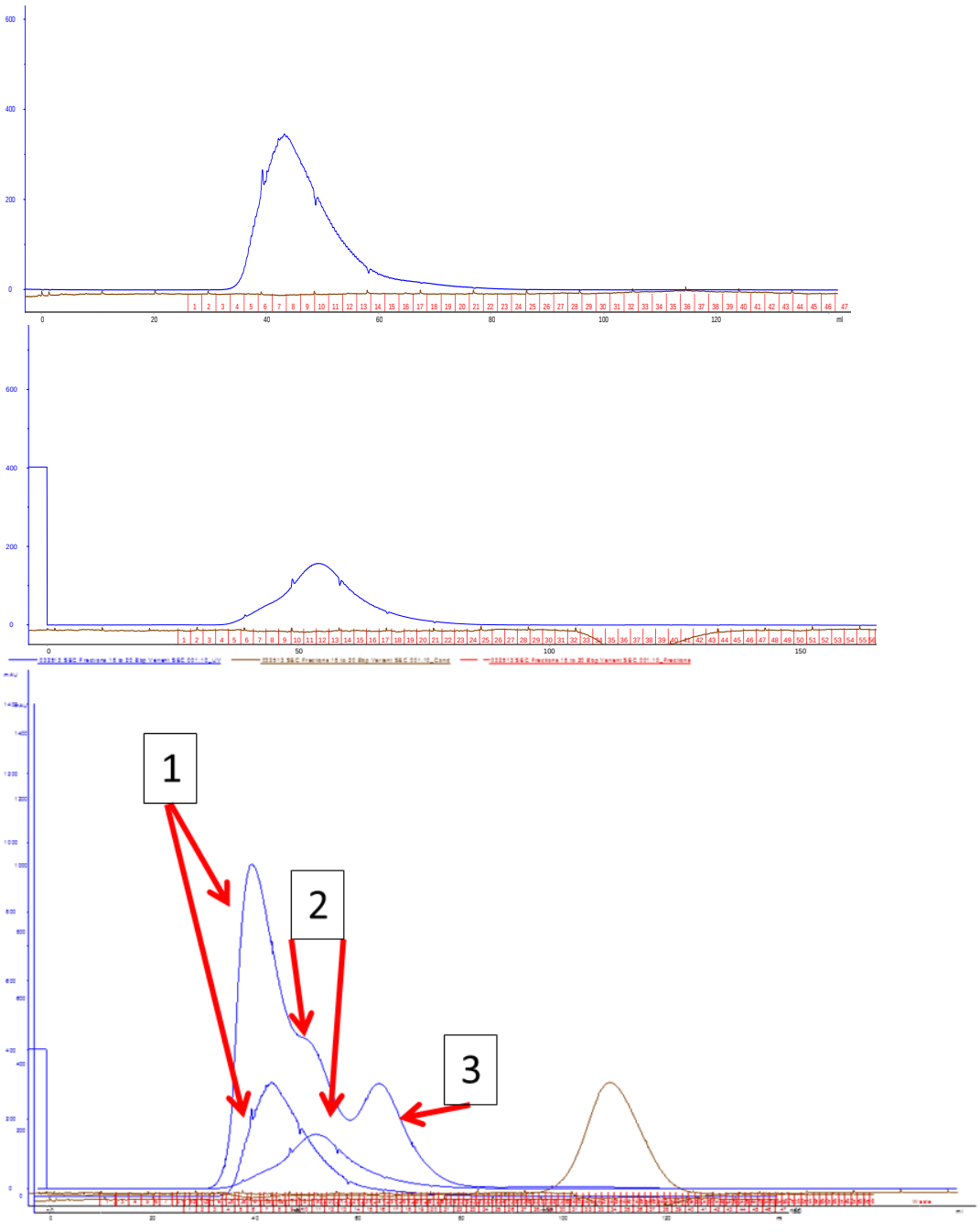


Figure 6-11. SdrE-Rydén SEC Repeat.
Panels 1 and 2 show the individual runs of the (1) multimer and (2) dimer peaks. As seen in the overlay (3) of these two runs with the original run, there is no change in the distribution from the (1) multimer peak fractions or (2) dimer peak fractions.

Upon a second SEC run after dilution and slow concentration, the data show that there was no change in the rate at which the multimer fractions or dimer fractions migrated through the column (Figure 6-11). While this does not disprove the hypothesis that the multimerization is a reversible event, it does suggest that any reversion to the monomeric state is likely to occur very slowly without the aid of secondary agents. However, it is useful to know that the monomer, dimer and multimer remained in their respective states because it allowed for further experimentation to determine if this multimerization results in a change in binding phenotype. To test for a change in phenotype, ELISA-type assays were used in which microtiter wells were coated with human fibrinogen, bone sialoprotein or mouse fibrinogen.

The multimerization state of SdrE2-Rydén-N2N3 affected the binding profile to human fibrinogen, bone sialoprotein and the negative control, pig fibrinogen (Figure 6-12). SdrE2-N2N3 only showed an affinity for human fibrinogen at these concentrations. Monomeric SdrE2-Rydén-N2N3 showed the greatest affinity for human fibrinogen also, but also displayed the ability to bind bone sialoprotein and pig fibrinogen. The dimeric form showed an equal affinity for all three ligands while the multimeric form is notable for the apparent loss of affinity for bone sialoprotein. The apparent interaction with pig fibrinogen by SdrE2-Rydén-N2N3 in all three states raised the concern of non-specific interactions.

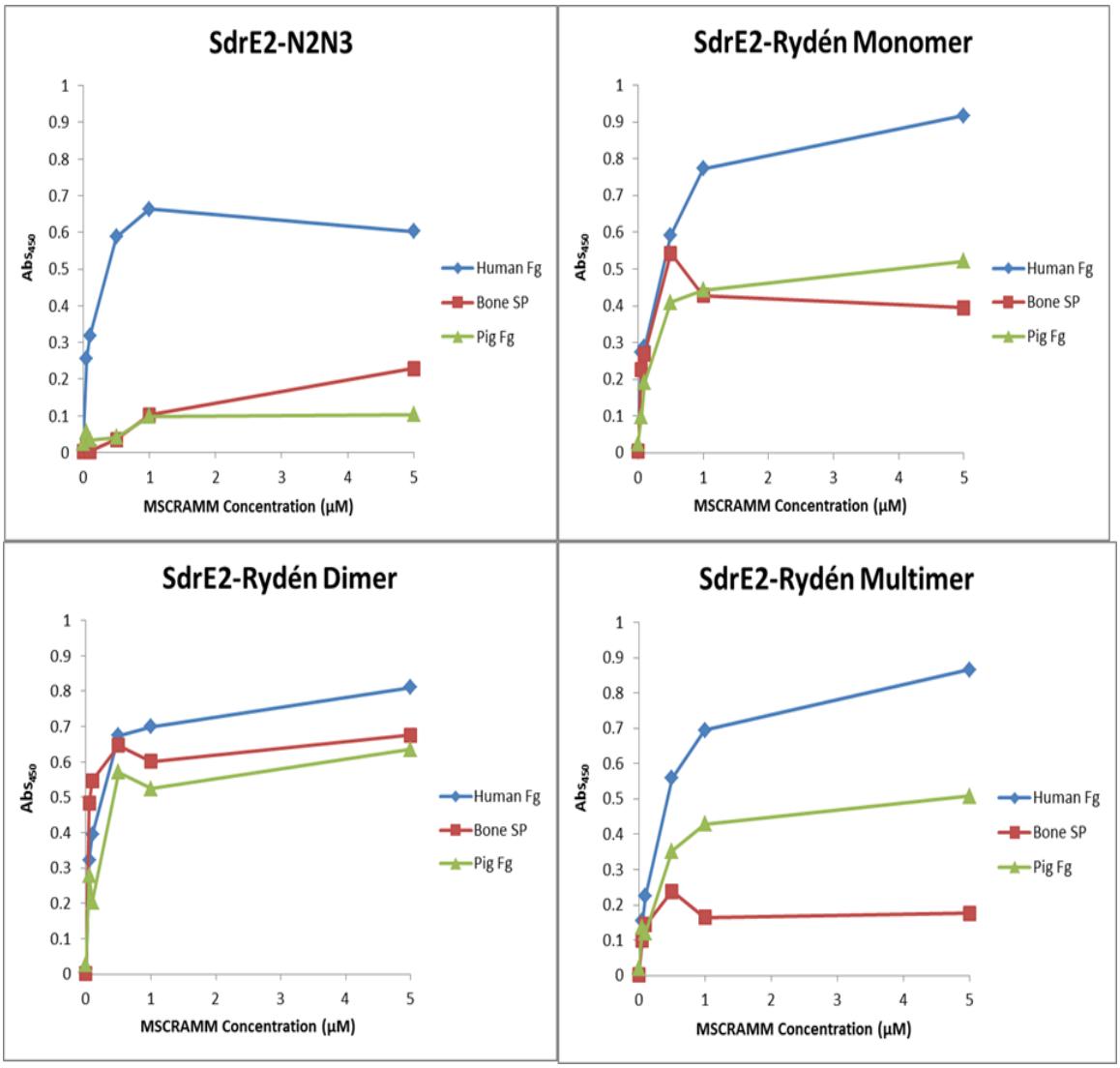


Figure 6-12. Fibrinogen and Bone Sialoprotein binding by SdrE-Rydén.

ELISA-type assays show different binding profiles for (1) wild type SdrE-N2N3, (2) SdrE-Rydén Monomer, (3) SdrE-Rydén Dimer, (4) SdrE-Rydén Multimer.

Peptide Inhibition ELISA-type assays were performed to confirm that SdrE2-Rydén-N2N3 interacts with human fibrinogen at the previously described target site and to test for binding of the BSP peptide or the BSP/Fg chimeric peptides (Figure 6-13). The lack of BSP binding in ELISA-type assay and the difficulty in measuring or estimating the number of SdrE2-Rydén-N2N3 molecules in each multimer lead us to focus on the monomeric and dimeric species in these studies.

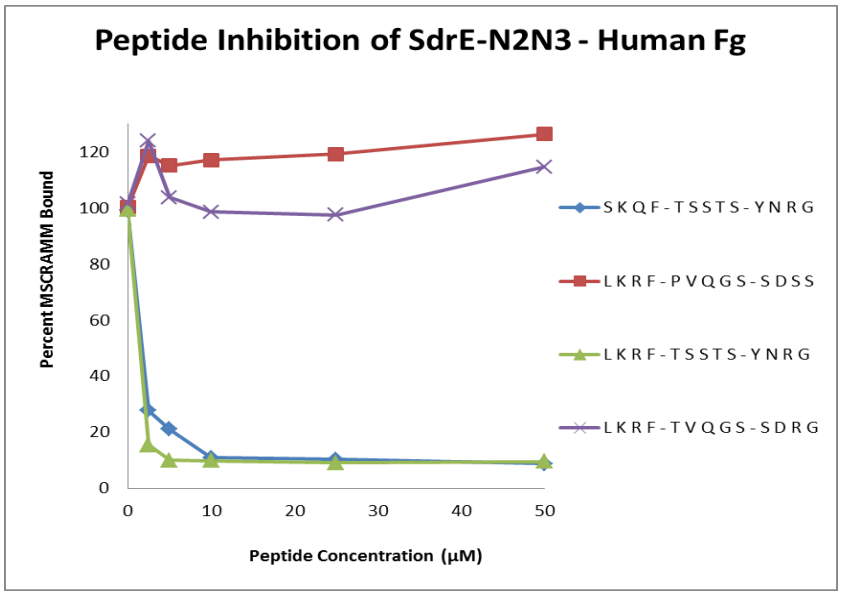
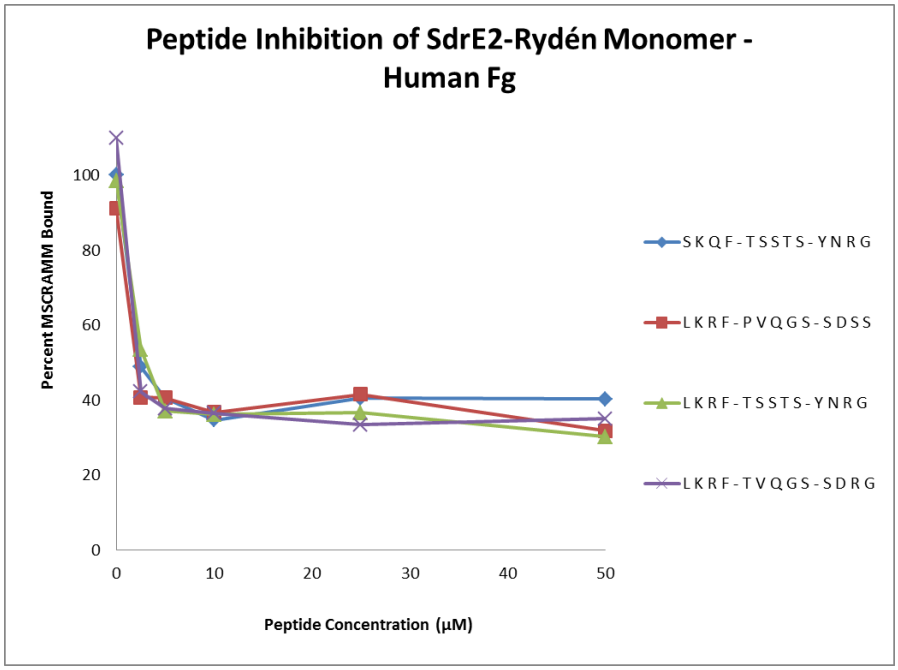


Figure 6-13. Peptide Inhibition of Wild Type SdrE-N2N3

The monomeric and dimeric forms of SdrE2-Rydén-N2N3 showed a significantly different peptide inhibition profile than that of SdrE2-N2N3. SdrE2-N2N3 shows a strong affinity for the wild type fibrinogen peptide and the BSP/Fg chimeric peptide in which the first four peptides are from BSP (Figure 6-14). This dose-dependent affinity results in 90% inhibition at a 1:100 ratio of MSCRAMM:Peptide (50μM). It shows no affinity for the BSP peptides that contain point mutations to the Fg sequence.

In contrast to SdrE, the monomeric SdrE2-Rydén-N2N3 shows 60% inhibition by all of the peptides at very low concentrations, which suggests a very high affinity for each of them. However, none of the peptides are able to achieve the 90% maximum inhibition as is seen with SdrE2-N2N3 (Figure 6-15). This is an unexpected finding because peptides that show 50% inhibition of MSCRAMMs binding their respective ligands in Peptide-inhibition ELISA-type assays at low concentrations are almost always able to fully inhibit the interaction at higher concentrations.



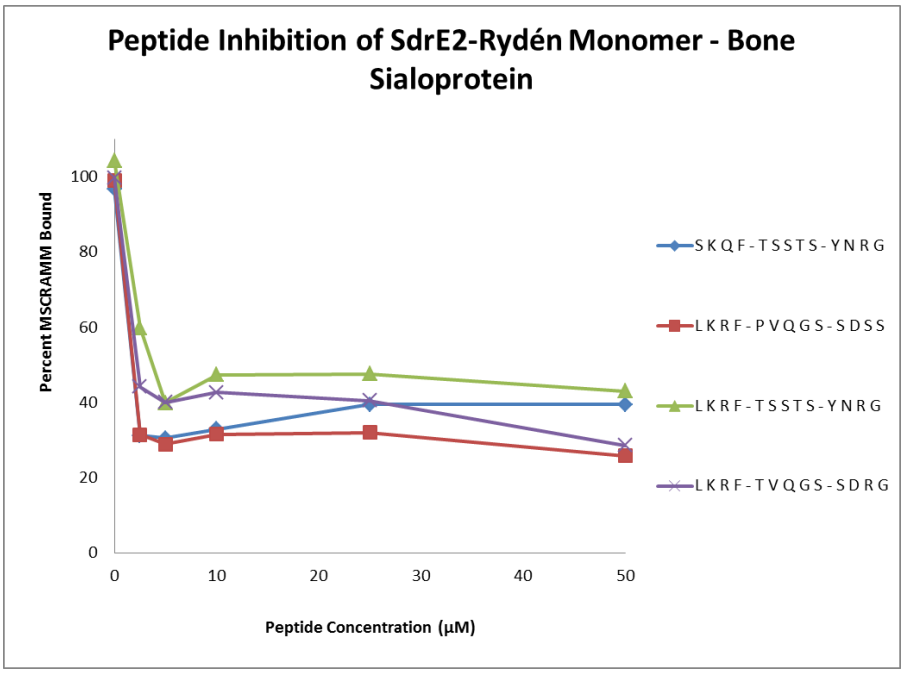
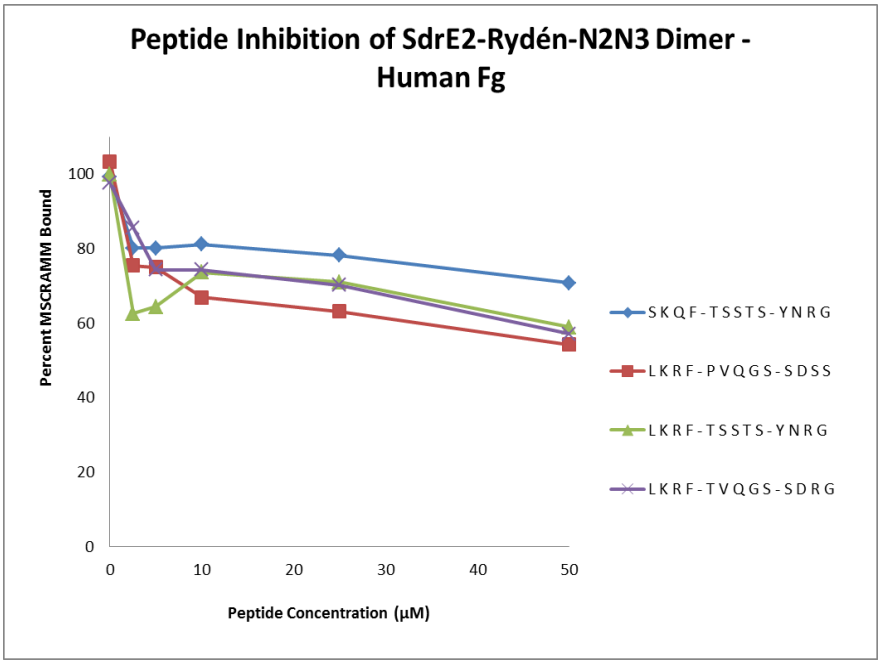


Figure 6-14. Peptide Inhibition of SdrE2-Rydén Monomer.
Panel 1 shows inhibition of SdrE2-Rydén monomer binding to fibrinogen. Panel 2 shows inhibition of SdrE2-Rydén monomer binding to bone sialoprotein.

Interestingly, the BSP peptide shows the best inhibition of monomeric SdrE2-Rydén-N2N3 binding to bone sialoprotein. These four peptides show a 20-40% inhibition of dimeric SdrE2-Rydén-N2N3 binding to human fibrinogen, but only one peptide shows a similar ability to inhibit binding to bone sialoprotein. This chimeric peptide is derived from the BSP peptide and contains two point mutations where the residues are changed to human fibrinogen. Given the lower levels of inhibition seen with the dimer as compared to the monomer, it appears that dimerization has affected the specificity of binding of SdrE2-Rydén-N2N3.

Attempts were made to study the interaction of the monomeric SdrE2-Rydén-N2N3 with the human Fg peptide, BSP petide and BSP/Fg chimera peptide in isothermal titration calorimetry experiments. However, no clear results were seen due to lack of consistently measurable heat given off during the binding reaction.



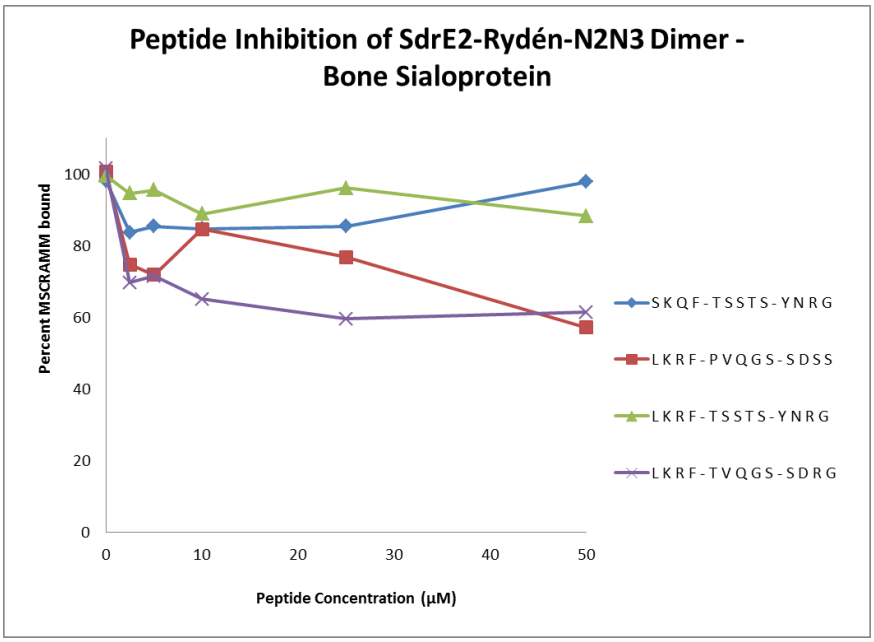


Figure 6-15. Peptide Inhibition of SdrE2-Rydén Dimer.
Panel 1 shows inhibition of SdrE2-Rydén dimer binding to fibrinogen. Panel 2 shows inhibition of SdrE2-Rydén dimer binding to bone sialoprotein.

DISCUSSION

Bbp/SdrE2 was originally named for its ability to bind to bone sialoprotein; however, whether the most commonly found version of SdrE2 is actually capable of binding BSP with high affinity remains to be proven. If Bbp/SdrE2 is unable to bind bone sialoprotein, then recent findings of is ability to bind to fibrinogen with high affinity and it being an allelic variant of SdrE1 strongly suggest that a change in nomenclature is in order. A change in nomenclature to SdrE2 would allow for greater clarity in the field when reporting the incidence, frequency and variations within these allelic variants.

Mutational analysis of the reported target sequence within BSP suggests that the first four residues, previously determined to be important for binding, are able to interact with SdrE2-N2N3. This indicates that the lack of binding to the reported BSP target sequence is due to one or multiple differences in the subsequent residues in the sequence.

The report of the SdrE2-Rydén sequence in 2000 raised the possibility that the previously published data regarding BSP binding involved an SdrE2 variant from an osteomyelitis isolate that had accumulated a significant number of mutations in the Lock-Latch region, shown above (Chapter 4) to be critical for binding. The SdrE2-Rydén-N2N3 construct was made through Overlap Extension PCR of the SdrE2-N2N3 construct and was then tested in ELISA-type assay and Peptide Inhibition ELISA-type assay.

The changes in sequence resulted in the generation of dimers and multimers during the expression and purification process. These multimers displayed a different binding profile from the monomer and had to be purified via size exclusion chromatography. We hypothesize that this is occurring because the introduction of a large number of hydrophobic residues to a region that is normally flexible, hydrophilic and exposed to the surrounding environment results. Multimerization may be induced by the hydrophobic effect where these residues are buried to prevent interaction with aqueous solvents.

It is unclear at this time whether SdrE2-Rydén-N2N3 represents a true virulence factor with a bone sialoprotein-binding mechanism of virulence or some form of error in cloning or sequencing. It should be noted that the nucleotide sequence is the same except for a nucleotide insertion right before the Lock region, a frameshift mutation that causes changes in all of the subsequent codons until a nucleotide deletion that occurs at the end of the Latch region resets the reading frame.

Interestingly, SdrE2-N2N3 did show an interaction with bone sialoprotein in ELISA-type assay, but it appeared to be a minor interaction of modest affinity when compared to the affinity of SdrE2-N2N3 for human fibrinogen. Future work is expected to confirm that while SdrE2-N2N3 can bone sialoprotein, this interaction is weak secondary form of virulence.

CHAPTER VII

SDRE1 BINDS FACTOR H

INTRODUCTION

Other ligands for SdrE

The observed frequencies of *sdrE1* and *sdrE2* in staphylococcal isolates gathered from non-human species is significantly different when compared to isolates taken from humans. *sdrE1* is present in 88% of staphylococcal isolates from a survey of animal species, while *sdrE2* is found in 11% of these isolates. In human disease, *sdrE1* is found in 55% and *sdrE2* is found in 32% of isolates. Interestingly, our lab has shown that the biochemical binding profiles of *sdrE1* and *sdrE2* provide a rationale for the observed overrepresentation of *sdrE1* in animal isolates. The ligand-binding N2N3 domain of SdrE2 shows a high affinity for human fibrinogen with a lesser ability to interact with bovine fibrinogen and no measurable binding capacity for fibrinogen from other species. In contrast, SdrE1-N2N3 shows a lesser affinity for human fibrinogen when compared to SdrE2-N2N3, but a more promiscuous binding profile that includes the ability to bind to fibrinogen from multiple species. These biochemical data provide a functional rationale for the observation that *S. aureus* colonizing a non-human host results in a genetic profile will skew towards a higher frequency of *sdrE1*$^{+}$*/sdrE2*$^{-}$ isolates.

However, when focusing on the gene frequencies of these allelic variants in human isolates only, the biochemical binding profiles of SdrE1-N2N3 and SdrE2-N2N3 do not adequately provide a rationale for the observed frequencies. SdrE2-N2N3 displays a 45-fold higher affinity for human fibrinogen than SdrE1-N2N3. If fibrinogen was the only host protein targeted by these allelic variants, *sdrE2* would be expected to be observed at significantly higher frequencies than *sdrE1*. One possible explanation for this observation is that *S. aureus* moves between species with enough frequency that retention of *sdrE1* imparts a competitive advantage. However, few published studies support host species transfer rates at such high levels.

Another, more likely, possibility is that the targeting of other human proteins by SdrE1 confer a competitive advantage. As mentioned previously, most MSCRAMMs show the ability to bind to multiple targets. The multifunctional capability of MSCRAMMs is important to bacteria that need to maintain efficient genomes. Given these previous findings regarding MSCRAMM targets, it would be expected that SdrE1 is capable of binding to multiple targets.

Furthermore, given that *sdrE1* occurs at greater frequencies in staphylococcal isolates taken from humans than *sdrE2,* we hypothesize that secondary targets of SdrE will be bound with higher affinity by *sdrE1* than *sdrE2*. In this model, *sdrE1* would provide a competitive advantage to *S. aureus* over *sdrE2* by allowing for multiple functionalities while *sdrE2* provides a single target, albeit one that is bound with greater affinity.

Factor H

In May, 2012, it was published that SdrE binds to Factor H, a member of the alternative complement pathway.[45] This group has published previously that another MSCRAMM in the Sdr family, ClfA, binds to Factor I, another member of the alternative complement pathway.[35] The complement system is an integral part of the innate immune defense whereby pathogens and foreign bodies are targeted and removed via opsonization and localized anaphylactic processes. While all three pathways of the complement system result in the generation of C3 convertases, in the alternative complement pathway, the C3 convertase C3bBb can be activated spontaneously. In order for the host to protect itself from non-specific activation and attack by the alternative pathway, soluble secreted inhibitors and cell membrane-bound inhibitors of this convertase are produced by the host. Factor I plays an inhibitory role in the pathway by enzymatically cleaving activated C3b to creat inactive iC3b. Factor H is another inhibitor of the alternative complement pathway that has two modes of action.[72] One mechanism involves initiating the irreversible disassociation of C3b and Bb. The second mechanism involves working as a co-factor for the enzymatic cleavage by Factor I of C3b.

Factor H in human disease

The importance of the role of Factor H in binding to human cells and inhibiting the alternative complement system is seen in a variety of pathologies that have been linked to mutations and single nucleotide polymorphisms in the Factor H gene. It has recently been reported that 35% of individuals are carriers of a Y402H mutation in Factor H; this mutation is correlated to an increased risk in age-related macular degeneration.[73] Homozygous and heterozygous carriers have sevenfold and threefold increased risks for the disease, respectively. Additionally, a subset of atypical hemolytic uremic syndrome is strongly linked to mutations in Factor H and other regulators of the complement system. Factor H mutations or deletions can also result in Membranoproliferative Glomerulonephritis type II (MPGN II) due to the resulting increased levels of circulating C3 activation products (C3a and C3b).[74]

Factor H structure

Factor H is a secreted 155 kDa serum protein. It is composed of 20 repeat regions, termed SCRs, and has an overall structure that resembles a beads-on-a-string appearance. The size and flexibility of Factor H has made crystallization of the whole protein difficult to achieve, but smaller portions of the structure have been elucidated. Analysis through use of truncation mutations has isolated the various functionalities and target domains of Factor H to different sets of SCRs.

Factor H is bound to host cells via its C-terminal SCRs, 18-20, which interact with the glycosaminoglycans that are attached to the surface of mammalian cells but are absent in bacterial and yeast cells. There are multiple sites for the binding of heparin and C3b. Recently, it was published that soluble Factor H dimerizes and that this dimerization occurs through SCRs 5-7 and 19-20.[75]

The potent inhibitory effect of Factor H on the alternative complement pathway is exploited by a variety of pathogens to evade the innate immune system. Yeast such as *Candida albicans,* the spirochete *Borrellia burgdorferi,* the Gram-negative *Neisseria gonorrhea,* and other Gram-positives such as *Streptococcus pyogenes* and *Streptococcus pneumonia* express proteins that bind to Factor H. These pathogens coat themselves with Factor H, resulting in local dissociation of the C3bBb convertase and increased inactivation of C3b to iC3b.[74]

RESULTS

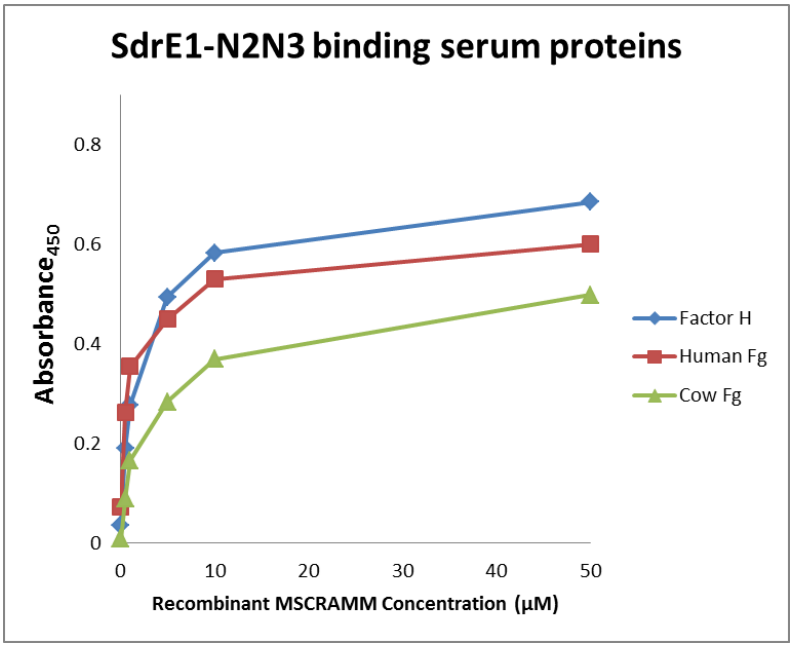
SdrE1-N2N3 binds Factor H in a specific manner with maximal binding in ELISA-type assay

In Sharp *et al*, the ELISA data did not show saturation in the interaction of SdrE1 with Factor H, thus making estimations of K_D impossible. This raises the possibility that the observed binding was a non-specific interaction. If SdrE1

interacts with Factor H in a non-specific manner or with low affinity, it is unlikely to be an effective mechanism of virulence.

In order to confirm that SdrE1 does bind Factor H, ELISA-type assays were performed using Factor H-coated wells and increasing amounts of purified recombinant SdrE1-N2N3. While the recombinant SdrE construct in Smith *et al* was composed of the SdrE A domain, the recombinant protein used in these experiments is the N2N3 constructs that were used previously to study the structure-function relationship of the SdrE-Fibrinogen interaction.

Also, in contrast to Sharp *et al*, where microtiter wells were coated with recombinant MSCRAMMs and then probed with increasing amounts of Factor H, these experiments reverse this process by probing Factor H-coated microtiter wells with recombinant MSCRAMMs. This is a more accurate reflection of the way in which these molecules interact in the *in vivo* environment.



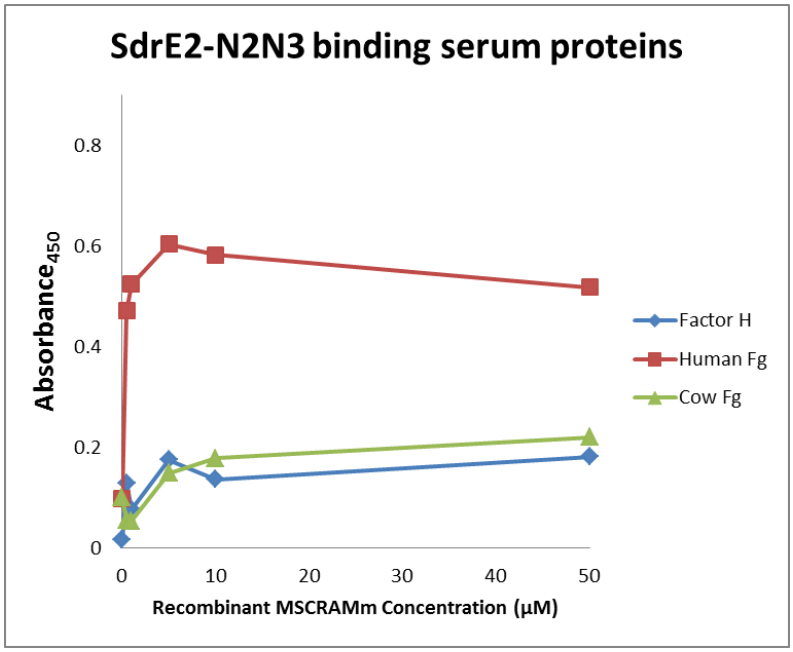


Figure 7-1. SdrE1-N2N3 binds Factor H, while SdrE2-N2N3 does not.
Panel 1 shows that SdrE1-N2N3 binds to Factor H to a similar degree as it does to Human Fibrinogen. Panel 2 shows that SdrE2-N2N3 does not appear to bind to Factor H.

Figure 7-1 shows that SdrE1-N2N3 binds to Factor H-coated microtiter wells with a similar affinity as SdrE1-N2N3 binds to human fibrinogen-coated microtiter wells. SdrE1 shows a promiscuous binding profile, as seen by its saturable, dose-dependent interaction with cow fibrinogen in this experiment. SdrE1 also binds Factor H strongly with an apparent K_D = 1 μM. Furthermore, SdrE2-N2N3, previously shown to be a highly specific, high affinity binder of human fibrinogen, does not interact with Factor H in these assays. Thus, Factor H provides an example of a ligand that satisfies our hypothesis regarding an alternate human target for SdrE in that SdrE1-N2N3 binds to it with significantly greater affinity than SdrE2-N2N3. Attempts were made to duplicate the Sharp *et al* data by probing MSCRAMM-coated wells with Factor H, however no binding was seen. This could be due to differences in affinity of primary anti-fH antibody. Alternatively, this could be explained by differences in SdrE affinity for Factor H based on whether it is in the soluble, flexible or coated, restrained state.

SPR analysis of SdrE allelic variants binding Factor H

Due to the size, flexibility and cost of Factor H, ITC experiments were not feasible. Furthermore, the inability to show binding via ELISA-type assay or SPR of soluble Factor H to bind to coated SdrE1-N2N3 suggests that it would be difficult to study this interaction in solution. Instead, SPR experiments were performed by Xiaowen Liang that analyzed the binding kinetics of the SdrE

allelic variants to immobilized Factor H. As shown in Figure 7-2, SdrE1-N2N3 binds tightly to Factor H , K_D = 0.16 μM. Surprisingly, SPR was also able to show an interaction between SdrE2-N2N3 and Factor H, but one with a 23-fold lower affinity when compared to SdrE1-N2N3 and Factor H.

We hypothesized that the differences in affinity for Factor H seen between the two SdrE allelic variants were due to the same structural features that were shown to be important in fibrinogen binding. To test this hypothesis, SdrE2-LockChimera was injected over the same Factor H-coated chip as SdrE1-N2N3 and SdrE2-N2N3 (Figure 7-3). The data show that SdrE2-LockChimera displayed a similar high affinity binding of Factor H that SdrE1-N2N3 showed (Table 7-1). This helps to confirm the hypothesis that the same structural features of SdrE are important for binding both fibrinogen and Factor H. It also lends further evidence that the Lock region plays an important role in ligand specificity for these allelic variants.

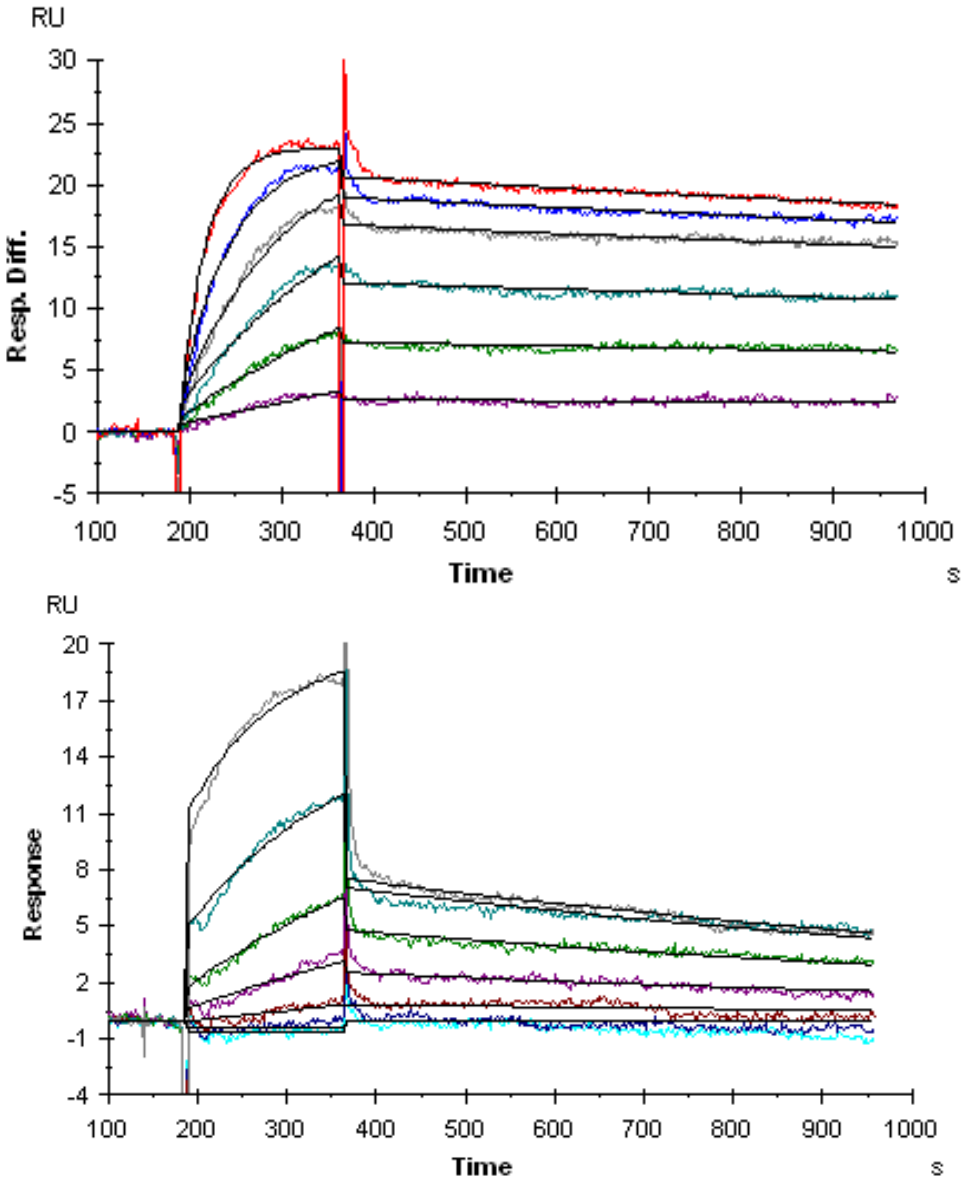


Figure 7-2. SPR measurement of SdrE1-N2N3 and SdrE2-N2N3 binding Factor H.

Panel 1 shows the promiscuous binding, SdrE1-N2N3 binding to Factor H in SPR experiments. Panel 2 shows SdrE2-N2N3 binding to Factor H at a significantly lower affinity.

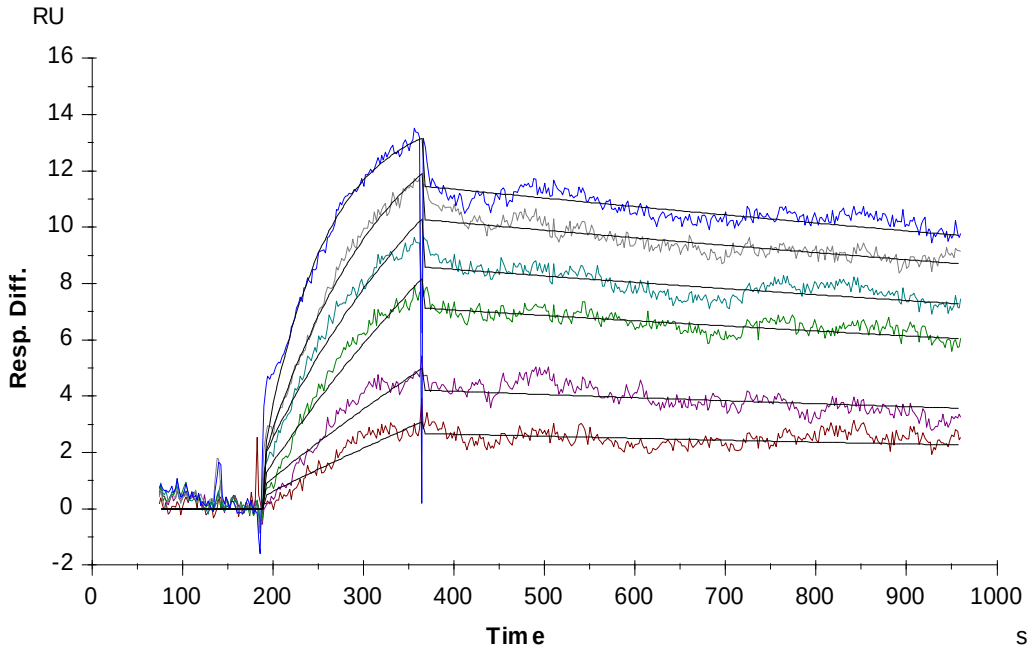


Figure 7-3. SdrE2-LockChimera binds Factor H similarly to SdrE1-N2N3.
The sensogram for SdrE2-LockChimera binding to Factor H.

Table 7-1. Kinetic measurements via SPR of SdrE binding Factor H

protein	k_a x10^3 ($M^{-1}s^{-1}$)	k_d x10^{-4} (s^{-1})	K_D (μM)	R_{max} (RU)
SdrE1-N2N3	1.14	1.90	0.16	22
SdrE2-N2N3	0.21	8.23	3.90	10
SdrE2-LockChimera	1.79	2.79	0.16	15

Progress towards isolating the target site of SdrE on Factor H

Attempts were made to find the specific target sequence in Factor H through two main approaches. First, the site in human fibrinogen that is targeted by SdrE1 and SdrE2, a 15mer from the fibrinogen Aα chain, was used to query for similar sequences in the Factor H amino acid sequence. The Factor H region that had the highest similarity was found to be Factor H 820-832, which occurs in SCR 13. Peptide inhibition ELISA-type assays were performed as described previously (Figure 7-4). While the human Fg peptide was able to inhibit the interaction between the SdrE allelic variants and fibrinogen, the Factor H peptide was not able to inhibit in a measurable concentration range.

In contrast to SdrE, the monomeric SdrE2-Rydén-N2N3 shows 60% inhibition by all of the peptides at very low concentrations, which suggests a very high affinity for each of them. However, none of the peptides are able to achieve the 90% maximum inhibition as is seen with SdrE2-N2N3 (Figure 6-15).

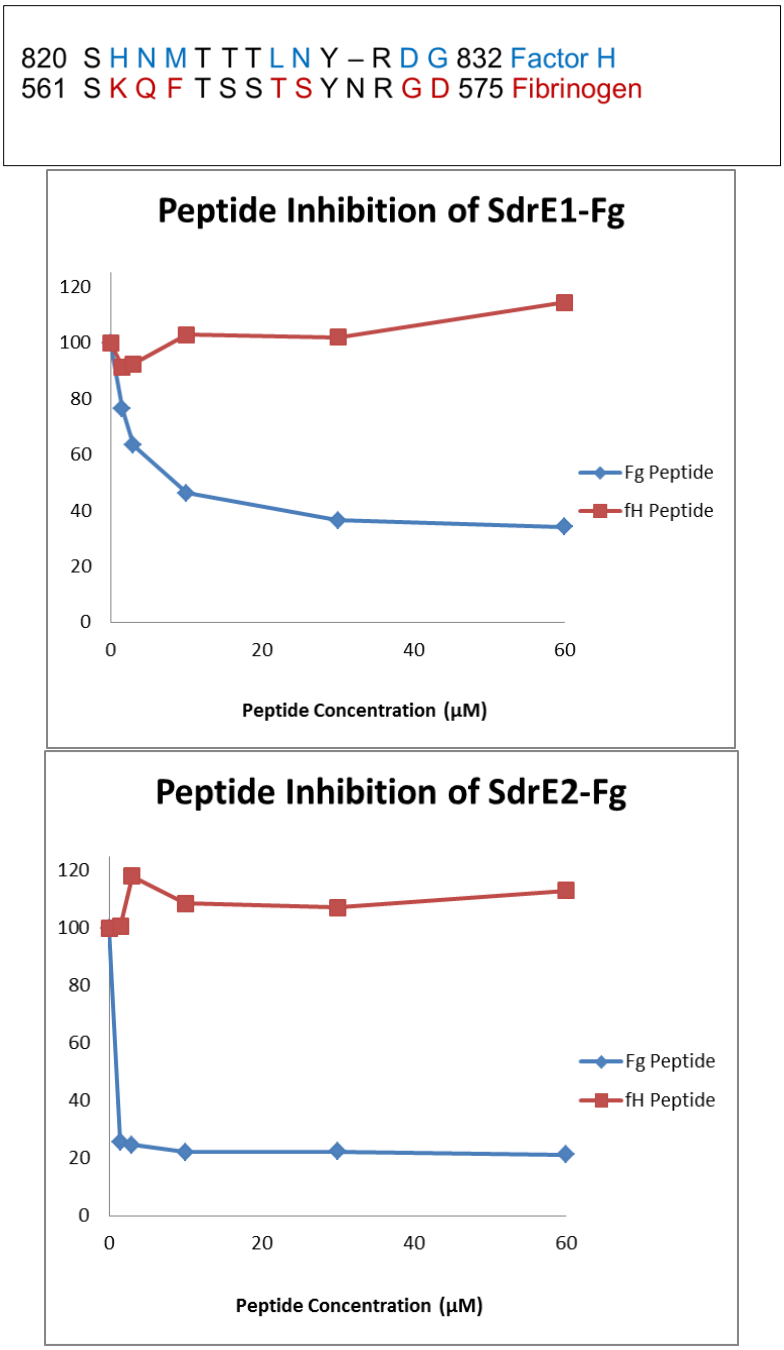


Figure 7-4. Putative Factor H binding site does not inhibit SdrE – Fibrinogen.

Panel 1 shows the putative Factor H binding site, designed from alignment studies with the human fibrinogen binding site. Panels 2 and 3 show Peptide Inhibition ELISA-type assays using these peptides to inhibit the binding of SdrE1-N2N3 and SdrE2-N2N3, respectively, to fibrinogen coated microtiter wells.

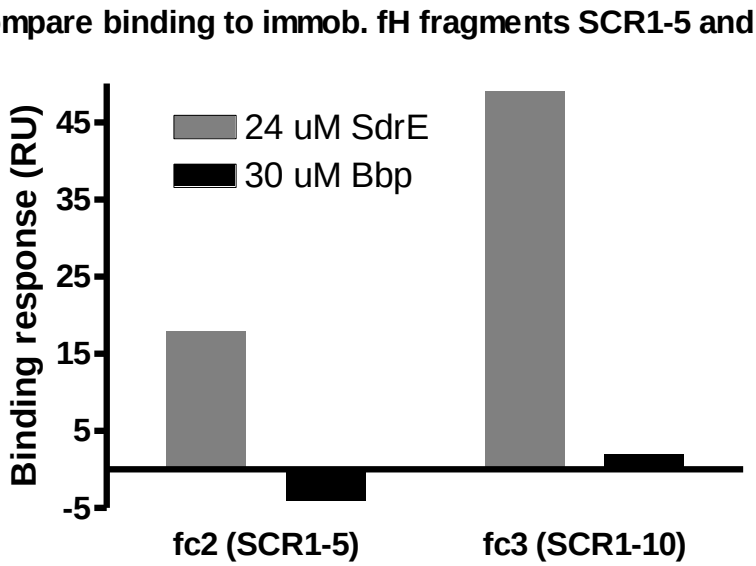


Figure 7-5. SdrE1 and Bbp/SdrE2 binding to Factor H fragments.

Our second approach to discover the binding site in Factor H used truncation mutants that were probed with the N2N3 domains of the SdrE allelic variants in SPR experiments performed by Xiaowen Liang (Figure 7-5). These data show that while Bbp/SdrE2-N2N3 is unable to bind to any fragments, SdrE-1 was able to bind to a truncation mutant containing the first ten SCRs. The inability of SdrE1-N2N3 to bind to the first 5 SCRs strongly suggests that the MSCRAMM is likely targeted to SCRs 5-10. Further experimentation narrows this range down to SCRs 5-7 (data not shown).

DISCUSSION

We have shown previously that the *in vitro* fibrinogen-binding characteristics of allelic variants SdrE1 and SdrE2 provide a rationale for the observed differences in gene frequencies between staphylococcal isolates gathered from humans and from animals. However, the human fibrinogen-binding profiles of these variants suggest a significantly different set of allelic frequencies should be expected in human staphylococcal isolates than what is observed. One hypothesis for this discrepancy is the existence of other host ligands provides a countervailing function for SdrE1. A recent publication provides evidence that SdrE1 binds Factor H, presenting a putative secondary ligand to match our hypothesis.

Building off of this initial report, we provide additional evidence that ligand-binding domain of SdrE1 binds to coated Factor H. While Sharp *et al* show an interaction, here we show data from analytical biochemical techniques that not only confirm the interaction, but provide information about the parameters of the binding event. Further work with coated Factor H truncation mutants will allow for isolation of the specific residues targeted by SdrE. It is possible that the binding of Factor H by SdrE1 does not follow the Dock, Lock, and Latch model as seen in the SdrE-Fibrinogen interaction. Factor H has a "beads on a string" conformation that only displays 4-8 amino acids on the "string" in between the SCRs. This makes it difficult to envision SdrE binding to a linear 11mer as seen with fibrinogen binding. If SdrE binds via a different

mechanism, then the search for a linear peptide of 10-15 amino acids is unlikely to yield conclusive results. Factor H truncation mutants representing single SCR domains, each approximately 60 amino acids, provides the best experimental approach for determining the binding site in this model. Alternatively, it is possible that coating of Factor H results in a conformational stretching that uncovers a linear sequence that is long enough for SdrE1 to attach via the Dock, Lock and Latch model.

Based on the high affinity displayed by SdrE1 for Factor H in the data presented above, we hypothesize that the SdrE1-Factor H interaction observed *in vitro* is important for *in vivo* virulence. This virulence mechanism is likely due to putative ability of Factor H-coated bacteria to inhibit the alternative complement pathway. Sharp *et al* provides evidence that presence of SdrE on the avirulent Gram positive bacteria *Lactococcus lactis* reduces the deposition of C3 fragments and C5a generation. Further experimentation is needed to confirm that full length SdrE1 on the surface of bacteria, specifically *S. aureus,* results in a similar reduction in C3 fragment deposition and complement fragment anaphylotoxins, as well as increased virulence *in vivo*. This could be done with *S. aureus* Newman Bald derivative strain, which has many MSCRAMMs removed, allowing for a clearer understanding of the role of SdrE1 in staphylococcal virulence.

It is possible that the higher affinity for Factor H of SdrE1 compared to SdrE2 is necessary and sufficient to explain the observed difference in gene

frequencies of these allelic variants in staphylococcal isolates taken from humans. However, the ability of MSCRAMMs to bind multiple targets suggests that the SdrE allelic variants could also target other host proteins. It is not clear if the ability of SdrE1 to bind Factor H provides an evolutionary advantage to *S. aureus* that overcomes the benefit of a higher affinity interaction with fibrinogen that SdrE2 provides. If there are additional ligands for these allelic variants involved, the respective affinities of SdrE1 and SdrE2 for new ligands would be difficult to predict based solely on epidemiological evidence. However, the biochemical data that we have gathered regarding the highly specific nature of SdrE2 binding to human fibrinogen suggests that SdrE2 will have low to no affinity for other ligands. The more promiscuous SdrE1would be the variant expected to be able to interact with other ligands.

In silico analysis was used to identify putative ligands for SdrE. A BLAST search was performed using the SdrE target sequence in human fibrinogen. A sequence from cytokeratin 14 was discovered from this search that is highly similar to the fibrinogen sequence (Figure 7-6). The ability of another MSCRAMM from the Sdr subfamily, ClfB, to interact with cytokeratin 10 has been published previously. This interaction is important for the ability of *S. aureus* to colonize the nares of the human host. SdrE binding of cytokeratin 14 could provide a similar mechanism for attachment and colonization. [37]

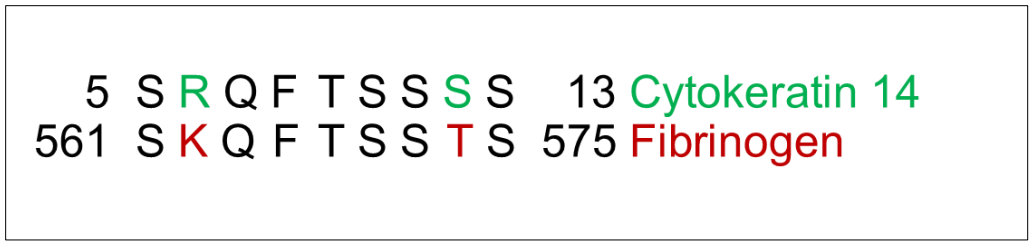


Figure 7-6. Cytokeratin 14 as a putative SdrE ligand.

It is interesting to note that the similarity occurs in the first nine amino acids of the human fibrinogen sequence. The solved crystal structure of SdrE1 in complex with this peptide shows contact points between these nine amino acids, while structural data from the SdrE2-Fibrinogen interaction shows that SdrE2 interacts with 13 amino acids in fibrinogen. Based on these data, it is predicted that if cytokeratin is bound by SdrE, SdrE1 is likely to bind with higher affinity than SdrE2.

CHAPTER VIII

CONCLUSION AND FUTURE DIRECTIONS

CONCLUSION

SdrE/SdrE1 and Bbp/SdrE2 are allelic variants

SdrE and Bbp have heretofore been regarded by most of the *S. aureus* literature as individual, unrelated members of the Sdr family of staphylococcal MSCRAMMs. SdrE is named for its location in the *S. aureus* core genome as the last member of the *sdrCDE* gene cluster. Bbp was named for its functional capacity as a bone sialoprotein binding protein. While there were some reports suggesting SdrE played a role in virulence, invasiveness and platelet activation, SdrE had no proven function until a recent publication showing that SdrE binds to Factor H.

Our research here shows that these MSCRAMMs are actually allelic variants. Bbp and SdrE have highly similar amino acid sequences that show 87% sequence identity and 95% sequence similarity within the A domain and B repeats. Furthermore, both genes are located in the same position in the *S. aureus* genome at the end of the *sdrCDE* gene cluster. Additionally, *sdrE1* and *sdrE2* are found together in less than 1% of *S. aureus* isolates. When they have been found in the same isolate, one generally occurs on a mobile genetic element. A portion of these double positive isolates are likely false positives that

result in counting one of the allelic variants twice due to the difficulty inherent to constructing primers for sequences that are so similar. Mapping the amino acid sequences of SdrE using dendogram analysis shows two allelic variants that bear more similarity to each other than any other MSCRAMM of the Sdr family, yet still segregate completely by allele with no crossover or intermediate variant.

In light of these data, we propose a change in nomenclature to dispel the notion that these are separate virulence factors. Given that Bbp/SdrE2 binds with high affinity to human fibrinogen and shows little affinity for bone sialoprotein, a name that reflects only bone sialoprotein binding is not ideal. In light of this and the higher frequency of SdrE, we propose a change in nomenclature of these allelic variants from SdrE to SdrE1 and Bbp to SdrE2.

When *sdrE1* and *sdrE2* genetic sequences are aligned, the number of variations can be looked at for each region of the gene. If *sdrE1* and *sdrE2* were individual virulence factors, it would be expected that the number of variations stayed constant across each domain. Instead, the N1 and B1-B3 domains are greater than 95% identical, while most of the variations are concentrated within the ligand binding N2 and N3 domains. In addition to providing additional evidence for the hypothesis that these two MSCRAMMs are actually allelic variants, there are a number of interesting applications for this observation. First, it is unusual that the domains that are currently predicted to only play a role in protrusion and display of the ligand binding domain are highly conserved while the functional, ligand binding domains are less conserved. This is the

opposite of the expected variation distribution, especially in comparison to critical residues in enzymes. It suggests that SdrE provides *S. aureus* with a somewhat redundant function resulting in less evolutionary pressure on the bacteria to maintain high fidelity during replication of these regions. The ability to allow for mutation in the effector domain of the virulence factor allows *S. aureus* to potentially adapt to its host by developing virulence factors that attach to new host proteins or to bind with greater affinity to its original target. Given that there are more variations within *sdrE1* alleles than *sdrE2* alleles, it seems likely that *sdrE2* came from *sdrE1*.

SdrE1 and *sdrE2* show significantly different distributions amongst isolates from humans and animals.

In *S. aureus* isolates taken from humans, *sdrE1* is found in approximately 55% of isolates while *sdrE2* is found in approximately 32% of isolates. In contrast, *S. aureus* isolates taken from a wide array of animals from multiple continents found that *sdrE1* is present in 88% of isolates while *sdrE2* is present in less than 11% of isolates. 75% of the isolates from animals that did contain *sdrE2* came from bovine species. These findings suggest that the distribution of *sdrE* allelic variants provides an example of *S. aureus* adaptation to different host species.

It is known that one mechanism of host adaptation displayed by *S. aureus* that has transitioned hosts from humans to animals is the rapid loss of virulence factors that are only functional against human targets. This phenomenon has been shown for staphylococcal virulence factors such as *clfA* and *hlaB*. The observed allelic distribution suggests that *sdrE1* and *sdrE2* provide another example of host adaptation with the loss of *sdrE2* in most isolates form animals except bovine species.

SdrE1 and SdrE2 display significantly different fibrinogen binding profiles

Our lab recently published that Bbp/SdrE2 binds to human fibrinogen Aα chain at residues 561-575.[47] Initial screens of the ligand binding N2N3 domain of SdrE2 with fibrinogen from other animals showed that SdrE2 displayed a high degree of specificity for human fibrinogen. Given the high degree of sequence identity between SdrE1 and SdrE2, we hypothesized that SdrE1 would also bind to human fibrinogen.

We show through ELISA-type assays, ITC and SPR experiments that SdrE1-N2N3 does bind to fibrinogen, but displays a significantly differently binding profile as compared to SdrE2-N2N3. SdrE2-N2N3 shows a 45-fold greater affinity for human fibrinogen than SdrE1-N2N3; however, SdrE1-N2N3 displays the ability to bind to fibrinogen from a broad sampling of species while

SdrE2-N2N3 only shows affinity for human fibrinogen and, to a much lesser extent, bovine fibrinogen.

Our molecular epidemiology data on SdrE strongly suggests that SdrE2 came from SdrE1. When viewed in combination with the biochemical data, it appears likely that *sdrE2* is an example of adaptation to the human host. The accumulated mutations in the N2N3 domain resulted in a significantly greater ability to bind to human fibrinogen, but also resulted in a reduced ability to bind to animal fibrinogen. Given that *S. aureus* is still mainly a human pathogen, it's not surprising that *sdrE2* is consistently present in these isolates at higher levels than what is seen in isolates from animals.

Interestingly, the fibrinogen binding profiles of SdrE1 and SdrE2 as measured in these biochemical experiments correlates strongly to the differences in allelic frequencies observed in staphylococcal isolates taken from animals compared to humans. SdrE1-N2N3 displays a broad species binding profile and the *sdrE1* gene is found at a higher frequency in isolates from animals when compared to the frequency of *sdrE1* in isolates taken from humans. Similarly, SdrE2-N2N3 displays little affinity for fibrinogen from most non-human species and *sdrE2* is found to have a significantly lower gene frequency in isolates taken from animals. Additionally, 9 of the 12 isolates that came from animals that were shown to contain *sdrE1* came from bovine species, while bovine fibrinogen was the only non-human fibrinogen that SdrE2-N2N3 bound to in biochemical assays.

With the growing incidence of Livestock Acquired-Methicillin Resistant *Staphylococcus aureus* (LA-MRSA) infections, there is a need to understand how a bacterium like *S. aureus* that is most often associated with human colonization and pathogenesis is able to move between such a large number of different species. These findings show that SdrE plays a role in *S. aureus* species adaptation with the presence of *sdrE1* generally conferring an advantage to *S. aureus* over *sdrE2*.

Understanding the structure-function relationship of the SdrE – Fg interaction

The significantly different fibrinogen binding profiles of *sdrE1* and *sdrE2* are surprising given the high degree of sequence identity between the two allelic variants within their ligand binding domains. In order to understand the structural basis for this biochemical observation, we collaborated with a crystallographer Dr. Ganesh Vannakambadi. This collaboration resulted in the solved crystal structures of SdrE2-N2N3 in an open conformation, as well as SdrE2-N2N3 in complex with the fibrinogen peptide representing the target site. These data support that SdrE2 binds to fibrinogen via the Dock, Lock and Latch mechanism.

Mutational analysis of the fibrinogen target sequence was performed in order to understand how SdrE2 displays a high specificity for human fibrinogen,

and to a lesser extent, bovine fibrinogen, given that there are so few differences within the corresponding target regions of fibrinogen from other species. Our data show that the valine insertion at position 565, which contains a threonine in humans, is capable reducing the affinity of SdrE2-N2N3 for this region by over 100-fold, as measured in Peptide Inhibition ELISA-type Assay with a peptide that contains a T565V mutation. Other changes, such as the subsequent R573G change, also affect binding, but do so at a much lesser level. As more human genome sequences become publicly available, it will be interesting to see if any mutations occur within this region and if those mutations result in a greater resistance to *S. aureus* colonization or infection.

Additionally, mutational analysis was performed to create a chimeric SdrE2-N2N3 that would display the binding phenotype of SdrE1-N2N3 while changing as few residues as possible. While no single point mutation resulted in an SdrE1-N2N3 binding phenotype, it was discovered that replacing the Lock domain of SdrE2 with the Lock domain from SdrE1 resulted in a chimeric construct did. SdrE2-LockChimera has 10 amino acids from SdrE1 yet displays a lower affinity for human fibrinogen and the ability to interact with fibrinogen from multiple species in a manner that is more similar to SdrE1 than SdrE2. The dramatic impact that such a small region can have on binding suggests that the Lock domains in SdrE and other MSCRAMMs should be monitored to watch for new variants that have greater virulence.

Factor H provides another human target and a rationale for observed allelic frequencies

The *in vitro* fibrinogen binding profiles of SdrE1 and SdrE2 provide a rationale for the observed differences in allelic distribution between staphylococcal isolates gathered from humans and animals. However, when restricting the focus to solely the allelic frequencies in human isolates, the binding profiles of SdrE1-N2N3 and SdrE2-N2N3 do not adequately provide a rationale for the observed allelic frequencies. SdrE2-N2N3 displays a 45-fold higher affinity for human fibrinogen than SdrE1-N2N3. If fibrinogen was the only host protein targeted by these allelic variants, *sdrE2* would be expected to be observed at significantly higher frequencies than *sdrE1*. Given that most MSCRAMMs bind multiple ligands, we hypothesized that there is a secondary human host protein for SdrE that SdrE1 binds with higher affinity than SdrE2.

A recently published study showed an interaction between SdrE1 and Factor H, an important regulator of the alternative complement pathway.[20] After showing that SdrE1 binds Factor H with a high affinity that reflects what is likely a biologically relevant interaction and play a role in virulence, we were able to show that Factor H satisfies the hypothesis of a secondary ligand that SdrE1 binds with significantly higher affinity than SdrE2. The original study published data that suggested the bound Factor H was functional and allowed for the local inhibition of the alternative complement pathway.

The many ways that *S. aureus* is able to inhibit and interfere with the complement system has been a virulence mechanism of growing interest over the past decade in our lab and others. While the importance of Factor H binding in comparison to Fg binding is still uncertain, we hypothesize that Factor H binding allows for greater virulence in the human host. This second function provides a rationale for the observation that *sdrE1* is found more frequently than *sdrE2* in staphylococcal isolates from humans.

Fibrinogen binding is likely more important than BSP binding

Bbp/SdrE2 was originally named for its ability to bind to bone sialoprotein. These original reports also found no fibrinogen binding. In contrast, our lab published that Bbp/SdrE2 binds to fibrinogen but has had difficulty to show an interaction of this MSCRAMM with BSP. We were able to acquire BSP purified from a human cell line, which had not previously been available to our lab and is important because it allows for the full post-translational modification of the gene product.

The ability of Bbp/SdrE2 to bind to BSP was weak as measured in ELISA-type assay. Given these data and the fibrinogen binding data, it is unlikely that BSP-binding represents the major function of Bbp/SdrE2. That fibrinogen binding is likely to be the more important mechanism of virulence lends

additional justification to a change in nomenclature to more adequately reflect the allelic nature of Bbp/SdrE2 and the similarities to SdrE1.

PUTATIVE STAPHYLOCOCCAL VACCINE TARGET FOR ANIMALS

It is clear that one of the underlying causes of the spread of antibiotic resistance amongst *S. aureus* isolates is the administration of antibiotics to livestock. European countries and the United States have started to take steps to limit the types of antibiotics that can be given to these animals, but this creates a hardship for farmers and veterinarians who need ways to combat bacteria and the variety of infections that they cause. Indeed, organizations representing farmers in the United States have already pushed back against the restrictions on antibiotic use, specifically cephalosporin use, that the FDA attempted to impose in 2007. After that failed attempt, the FDA is again attempting to put restrictions in place, albeit in a more limited manner.[76]

An ideal way to combat these bacterial infections without the use of antibiotics would be a vaccine against *S. aureus* that could be used in animals. Ideal targets for this vaccine would be surface proteins, such as MSCRAMMs, that are shown to remain in the genome after the series of host adaptation steps occur. Although ClfA is an excellent putative target for a multivalent *S. aureus* vaccine in humans, the accumulation of truncation mutations in *clfA* genes from

LA-MRSA strains suggest that it is not expressed and displayed on the surface of LA-MRSA that has adapted to the animal host.

Our work here shows that SdrE, could be an excellent target for a vaccine against LA-MRSA to be given to animals. *sdrE1* is found in *S. aureus* isolates taken from an array of animals from multiple continents and shows a functionality in *in vitro* experiments that strongly suggests a role in virulence in animal hosts. SdrE1 has already been identified as a putative vaccine target in humans and the generation of antibodies against both SdrE1 and SdrE2 has been published in the literature. One possible concern is potential for mutation in *sdrE1* resulting in the generation of new variants by the bacteria due to selection pressures, but monitoring of the gene should allow a partial defense against this. *sdrE1* has shown greater variability than *sdrE2,* but the sequence identity is still very high between genes. It is possible that the structural data discussed here could be used to identify regions with high degrees of similarity between SdrE1 and SdrE2 that could be used as the basis of a vaccine effective against both variants.

OTHER IMPLICATIONS FOR FUTURE RESEARCH

The finding that the Lock domain of SdrE1/SdrE2 plays a large role in specificity and affinity of fibrinogen binding raises a number of important questions for future research. It is possible that more virulent *S. aureus* strains

evolve not only by the acquisition of a new gene or mobile genetic element, but also by accumulating mutations in small regions or subdomains that result in a significant alteration in the ligand binding profile. As greater understanding is gained of the domains within staphylococcal virulence factors that are critical for the binding phenotype, the need to monitor these shorts sequences will grow. Currently, most of the monitoring of sequence variations involves studying the sequences of antibiotic resistance genes. In the case of the SCC*mec* element, sequence variations in antibiotic resistance genes and recombinases are used for classification. Our data suggest that adhesins such as MSCRAMMs represent another important set of virulence factors that should be monitored in order to observe further host adaptation.

One potential example of continued mutation in this region is SdrE1-CC398. This clonal complex has been shown to play a major role in Livestock Acquired- MRSA infection. Given that SdrE2-N2N3 does not interact with fibrinogen from animals except for cows, it is not surprising that *sdrE1* is overrepresented in animal staphylococcal isolates and is the allelic variant found in *S. aureus* CC398. However, it is possible that the variations within SdrE1-CC398 provide an example of host adaptation to non-human species. The SdrE1-CC398 amino acid sequence contains a T587N mutation that occurs within the Lock domain. Furthermore, over 50% of the other variations found in the ligand binding domain of SdrE1-CC398 are changes that result in the corresponding amino acid from SdrE2. These findings suggest that SdrE-CC398 has

accumulated mutations that result in an altered ligand binding profile, potentially including a greater capacity for binding fibrinogen from animals. While there are no changes within the SdrE-CC398 sequence previously established to affect binding, the changes can be mapped onto the SdrE1-N2N3-Fg Peptide structure to provide greater understanding of how these variations could potentially affect the fibrinogen binding from animals.

Another implication of this research regards the rapidly increasing pace at which large sequencing projects are being started. New, powerful sequencing techniques are being used to look at many isolates at once and perform comparative genomic analyses. As sequencing capabilities make large scale sequencing projects on the order of 10^3 genomes possible, there are heretofore unanswered questions regarding how to adequately and appropriately to analyze genomic data that is orders of magnitude larger than what was available previously. Studies that have been published to date using these approaches have focused on data points such as the presence, lack of or movement of large mobile genetic elements. Examples include φSa3, which has been found in staphylococcal isolates from avian strains and contains genes that have been putatively identified as virulence factors for avian species based on similarity studies. Furthermore, these studies have noted the accumulation of stop codons or disappearance altogether of genes like *clfA, clfB,* and *hlaB*.

However, sequencing data alone cannot provide the rationale for the observed differences in allelic frequency in *sdrE* in animal and human

staphylococcal isolates. While all of the residues that play a role in the differing phenotype of SdrE1 and SdrE2 have yet to be elucidated, our data clearly show that the Lock domain is critically important for binding and that variations within this region can result in phenotypic differences. The sequence changes between the variants were not changes that would clearly result in a change in phenotype based on *in silico* analysis. However, through careful biochemical experiments and structural data, this domain was elucidated. In this instance, the biochemical data has provided information to explain the sequencing data. And now, this information can be used to more capably interpret future sequencing data.

SdrE is not the only staphylococcal virulence factor that has been shown to accumulate mutations that result in a change in binding profile. In staphylococcal infections of implanted cardiovascular devices, FnBPA accumulates polymorphisms that are responsible for a greater affinity for fibronectin as measured by atomic force microscopy.[30] This provides another example of host adaptation by *S. aureus*, albeit over a shorter time frame and on a smaller scale than seen with the CC398 studies.

Similar to allelic variants SdrE1 and SdrE2, SCIN proteins are staphylococcal complement inhibitors that have three active members in the genome. Despite a similar overall structure and targeting of the same spots on C3 convertase, there were key differences between SCIN-A and SCIN-B in specific residues that bound C3 convertase and the manner in which they bound.

Structural data revealed the basis for a lack of SCIN-D binding. Together, these examples speak to the need to understand the biochemical changes that sequence variations can cause at the molecular level.

To gain a better understanding of the effects of sequence variations that are found in large scale whole genome sequencing projects, biochemical approaches will be necessary. However, many of the approaches used here require large amounts of highly purified protein and a significant amount of time. In order to keep pace with the rapidly expanding database of whole genome sequences and the variations in virulence factors that will be discovered, there will be a greater need for high throughput biochemical approaches to examine the phenotypic effects of small variations and mutations.

It is facile to think of *Staphylococcus aureus* as having a static, if diverse, set of virulence factors that allow for classification and eventual targeting by antibacterial or vaccine approaches. *S. aureus* should instead be seen as an rapidly evolving bacteria that has had decades, if not longer, to adapt to the colonization and infection of the human host. In the past, the absence, presence and emergence of new genes were the focus of studies of *S. aureus* virulence. However, the rapidly growing capabilities of sequencing technology will result not just in a larger data set of whole genome sequences, but greater potential to pinpoint small sequence changes that result in significant changes in binding or virulence. If high throughput biochemical techniques are developed to pair with the large data sets of whole genome sequences, we will gain far greater

understanding of both the mechanisms of virulence of *S. aureus* and the ways in which it is currently adapting and evolving to challenge the defenses of the non-human and human hosts.

REFERENCES

1. Dworkin M, Götz, FG, Bannerman T & Schleifer K. *The Genera Staphylococcus and Macrococcus*. New York, N.Y.: Springer; 2006.

2. Lowy FD. Staphylococcus aureus infections. *N Engl J Med*. Aug 1998;339(8):520-532.

3. Otto M. MRSA virulence and spread. *Cell Microbiol*. Oct 2012;14(10):1513-1521.

4. Mediavilla JR, Chen L, Mathema B, Kreiswirth BN. Global epidemiology of community-associated methicillin resistant Staphylococcus aureus (CA-MRSA). *Curr Opin Microbiol*. Oct 2012;15(5):588-595.

5. Klevens RM, Morrison MA, Nadle J, et al. Invasive methicillin-resistant Staphylococcus aureus infections in the United States. *JAMA*. Oct 2007;298(15):1763-1771.

6. Nichol KA, Adam HJ, Roscoe DL, et al. Changing epidemiology of methicillin-resistant Staphylococcus aureus in Canada. *J Antimicrob Chemother*. May 2013;68 Suppl 1:i47-i55.

7. Karampela I, Poulakou G, Dimopoulos G. Community acquired methicillin resistant Staphylococcus aureus pneumonia: an update for the emergency and intensive care physician. *Minerva Anestesiol*. Aug 2012;78(8):930-940.

8. Mruk AL, Record KE. Antimicrobial options in the treatment of adult staphylococcal bone and joint infections in an era of drug shortages. *Orthopedics*. May 2012;35(5):401-407.

9. Calhoun JH, Manring MM, Shirtliff M. Osteomyelitis of the long bones. *Semin Plast Surg*. May 2009;23(2):59-72.

10. Defres S, Marwick C, Nathwani D. MRSA as a cause of lung infection including airway infection, community-acquired pneumonia and hospital-acquired pneumonia. *Eur Respir J*. Dec 2009;34(6):1470-1476.

11. Balaban N, Rasooly A. Staphylococcal enterotoxins. *Int J Food Microbiol*. Oct 2000;61(1):1-10.

12. Le Loir Y, Baron F, Gautier M. Staphylococcus aureus and food poisoning. *Genet Mol Res*. 2003;2(1):63-76.

13. Dinges MM, Orwin PM, Schlievert PM. Exotoxins of Staphylococcus aureus. *Clin Microbiol Rev*. Jan 2000;13(1):16-34

14. Young AE, Thornton KL. Toxic shock syndrome in burns: diagnosis and management. *Arch Dis Child Educ Pract Ed*. Aug 2007;92(4):ep97-100.

15. Sousa C, Botelho C, Rodrigues D, Azeredo J, Oliveira R. Infective endocarditis in intravenous drug abusers: an update. *Eur J Clin Microbiol Infect Dis*. Nov 2012;31(11):2905-2910.

16. Westphal N, Plicht B, Naber C. Infective endocarditis--prophylaxis, diagnostic criteria, and treatment. *Dtsch Arztebl Int*. Jul 2009;106(28-29):481-489;

17. Chambers HF, Deleo FR. Waves of resistance: Staphylococcus aureus in the antibiotic era. *Nat Rev Microbiol*. Sep 2009;7(9):629-641.

18. Bayer AS, Schneider T, Sahl HG. Mechanisms of daptomycin resistance in Staphylococcus aureus: role of the cell membrane and cell wall. *Ann N Y Acad Sci*. Jan 2013;1277:139-158.

19. Kloos WE. Natural populations of the genus Staphylococcus. *Annu Rev Microbiol*. 1980;34:559-592.

20. Price LB, Stegger M, Hasman H, et al. Staphylococcus aureus CC398: host adaptation and emergence of methicillin resistance in livestock. *MBio*. 2012;3(1).

21. de Boer E, Zwartkruis-Nahuis JT, Wit B, et al. Prevalence of methicillin-resistant Staphylococcus aureus in meat. *Int J Food Microbiol*. Aug 2009;134(1-2):52-56.

22. Hanson BM, Dressler AE, Harper AL, et al. Prevalence of Staphylococcus aureus and methicillin-resistant Staphylococcus aureus (MRSA) on retail meat in Iowa. *J Infect Public Health*. Sep 2011;4(4):169-174.

23. Lamamy C, Berthelot A, Bertrand X, et al. CC9 livestock-associated Staphylococcus aureus emerges in bloodstream infections in French patients unconnected with animal farming. *Clin Infect Dis*. Apr 2013;56(8):e83-86.

24. McCarthy AJ, van Wamel W, Vandendriessche S, et al. Staphylococcus aureus CC398 clade associated with human-to-human transmission. *Appl Environ Microbiol.* Dec 2012;78(24):8845-8848.

25. Uhlemann AC, Porcella SF, Trivedi S, et al. Identification of a highly transmissible animal-independent Staphylococcus aureus ST398 clone with distinct genomic and cell adhesion properties. *MBio*. 2012; 3(2): e00027-12 .

26. Lowder BV, Guinane CM, Ben Zakour NL, et al. Recent human-to-poultry host jump, adaptation, and pandemic spread of Staphylococcus aureus. *Proc Natl Acad Sci U S A*. Nov 2009;106(46):19545-19550.

27. Cavaco LM, Hasman H, Aarestrup FM. Zinc resistance of Staphylococcus aureus of animal origin is strongly associated with methicillin resistance. *Vet Microbiol.* Jun 2011;150(3-4):344-348.

28. Mazmanian SK, Ton-That H, Schneewind O. Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus. *Mol Microbiol.* Jun 2001;40(5):1049-1057.

29. Rivera J, Vannakambadi G, Höök M, Speziale P. Fibrinogen-binding proteins of Gram-positive bacteria. *Thromb Haemost*. Sep 2007;98(3):503-511.

30. Casillas-Ituarte NN, Lower BH, Lamlertthon S, Fowler VG, Lower SK. Dissociation rate constants of human fibronectin binding to fibronectin-binding proteins on living Staphylococcus aureus isolated from clinical patients. *J Biol Chem.* Feb 2012;287(9):6693-6701.

31. Lower SK, Lamlertthon S, Casillas-Ituarte NN, et al. Polymorphisms in fibronectin binding protein A of Staphylococcus aureus are associated with infection of cardiovascular devices. *Proc Natl Acad Sci U S A*. Nov 2011;108(45):18372-18377.

32. McAdow M, Missiakas DM, Schneewind O. Staphylococcus aureus secretes coagulase and von Willebrand factor binding protein to modify the coagulation cascade and establish host infections. *J Innate Immun.* 2012;4(2):141-148.

33. Walsh EJ, Miajlovic H, Gorkun OV, Foster TJ. Identification of the Staphylococcus aureus MSCRAMM clumping factor B (ClfB) binding site in the alphaC-domain of human fibrinogen. *Microbiology*. Feb 2008;154(Pt 2):550-558.

34. Ganesh VK, Rivera JJ, Smeds E, et al. A structural model of the Staphylococcus aureus ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathog*. Nov 2008;4(11):e1000226.

35. Hair PS, Echague CG, Sholl AM, et al. Clumping factor A interaction with complement factor I increases C3b cleavage on the bacterial surface of Staphylococcus aureus and decreases complement-mediated phagocytosis. *Infect Immun*. Apr 2010;78(4):1717-1727.

36. Wertheim HF, Walsh E, Choudhurry R, et al. Key role for clumping factor B in Staphylococcus aureus nasal colonization of humans. *PLoS Med*. Jan 2008;5(1):e17.

37. Xiang H, Feng Y, Wang J, et al. Crystal structures reveal the multi-ligand binding mechanism of Staphylococcus aureus ClfB. *PLoS Pathog*. 2012;8(6):e1002751.

38. Ponnuraj K, Bowden MG, Davis S, et al. A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell*. Oct 2003;115(2):217-228.

39. Otto M. Molecular basis of Staphylococcus epidermidis infections. *Semin Immunopathol*. Mar 2012;34(2):201-214.

40. Sitkiewicz I, Babiak I, Hryniewicz W. Characterization of transcription within sdr region of Staphylococcus aureus. *Antonie Van Leeuwenhoek*. Feb 2011;99(2):409-416.

41. Arrecubieta C, Lee MH, Macey A, Foster TJ, Lowy FD. SdrF, a Staphylococcus epidermidis surface protein, binds type I collagen. *J Biol Chem*. Jun 2007;282(26):18767-18776.

42. Josefsson E, O'Connell D, Foster TJ, Durussel I, Cox JA. The binding of calcium to the B-repeat segment of SdrD, a cell surface protein of Staphylococcus aureus. *J Biol Chem*. Nov 1998;273(47):31145-31152.

43. Barbu EM, Ganesh VK, Gurusiddappa S, et al. beta-Neurexin is a ligand for the Staphylococcus aureus MSCRAMM SdrC. *PLoS Pathog*. Jan 2010;6(1):e1000726.

44. Corrigan RM, Miajlovic H, Foster TJ. Surface proteins that promote adherence of Staphylococcus aureus to human desquamated nasal epithelial cells. *BMC Microbiol*. 2009;9:22.

45. Sharp JA, Echague CG, Hair PS, et al. Staphylococcus aureus surface protein SdrE binds complement regulator Factor H as an immune evasion tactic. *PLoS One*. 2012;7(5):e38407.

46. Rydén C, Tung HS, Nikolaev V, Engström A, Oldberg A. Staphylococcus aureus causing osteomyelitis binds to a nonapeptide sequence in bone sialoprotein. *Biochem J*. Nov 1997;327 (Pt 3):825-829.

47. Vazquez V, Liang X, Horndahl JK, et al. Fibrinogen is a ligand for the Staphylococcus aureus microbial surface components recognizing adhesive matrix molecules (MSCRAMM) bone sialoprotein-binding protein (Bbp). *J Biol Chem*. Aug 2011;286(34):29797-29805.

48. Vengadesan K, Narayana SV. Structural biology of Gram-positive bacterial adhesins. *Protein Sci*. May 2011;20(5):759-772.

49. Ponnuraj K, Xu Y, Macon K, Moore D, Volanakis JE, Narayana SV. Structural analysis of engineered Bb fragment of complement factor B: insights into the activation mechanism of the alternative pathway C3-convertase. *Mol Cell*. Apr 2004;14(1):17-28.

50. DeLeo FR, Chambers HF. Reemergence of antibiotic-resistant Staphylococcus aureus in the genomics era. *J Clin Invest*. Sep 2009;119(9):2464-2474.

51. Li M, Du X, Villaruz AE, et al. MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med*. May 2012;18(5):816-819.

52. Rydén C, Yacoub AI, Maxe I, et al. Specific binding of bone sialoprotein to Staphylococcus aureus isolated from patients with osteomyelitis. *Eur J Biochem*. Sep 1989;184(2):331-336.

53. Yacoub A, Lindahl P, Rubin K, Wendel M, Heinegård D, Rydén C. Purification of a bone sialoprotein-binding protein from Staphylococcus aureus. *Eur J Biochem*. Jun 1994;222(3):919-925.

54. Otsuka T, Saito K, Dohmae S, et al. Key adhesin gene in community-acquired methicillin-resistant Staphylococcus aureus. *Biochem Biophys Res Commun*. Aug 2006;346(4):1234-1244.

55. Stephens AJ, Huygens F, Inman-Bamber J, et al. Methicillin-resistant Staphylococcus aureus genotyping using a small set of polymorphisms. *J Med Microbiol*. Jan 2006;55(Pt 1):43-51.

56. Peacock SJ, Moore CE, Justice A, et al. Virulent combinations of adhesin and toxin genes in natural populations of Staphylococcus aureus. *Infect Immun*. Sep 2002;70(9):4987-4996.

57. Montanaro L, Speziale P, Campoccia D, et al. Polymorphisms of agr locus correspond to distinct genetic patterns of virulence in Staphylococcus aureus clinical isolates from orthopedic implant infections. *J Biomed Mater Res A*. Sep 2010;94(3):825-832.

58. Monecke S, Luedicke C, Slickers P, Ehricht R. Molecular epidemiology of Staphylococcus aureus in asymptomatic carriers. *Eur J Clin Microbiol Infect Dis*. Sep 2009;28(9):1159-1165.

59. Sabat A, Melles DC, Martirosian G, Grundmann H, van Belkum A, Hryniewicz W. Distribution of the serine-aspartate repeat protein-encoding sdr genes among nasal-carriage and invasive Staphylococcus aureus strains. *J Clin Microbiol*. Mar 2006;44(3):1135-1138.

60. Kuhn G, Francioli P, Blanc DS. Evidence for clonal evolution among highly polymorphic genes in methicillin-resistant Staphylococcus aureus. *J Bacteriol*. Jan 2006;188(1):169-178.

61. Aronovich A, Tchorsh D, Shezen E, et al. Enhancement of pig embryonic implants in factor VIII KO mice: a novel role for the coagulation cascade in organ size control. *PLoS One*. 2009;4(12):e8362.

62. Tziomalos K, Vakalopoulou S, Perifanis V, Garipidou V. Treatment of congenital fibrinogen deficiency: overview and recent findings. *Vasc Health Risk Manag*. 2009;5:843-848.

63. Medved L, Weisel JW, Haemostasis FaFXSoSSCoISoTa. Recommendations for nomenclature on fibrinogen and fibrin. *J Thromb Haemost*. Feb 2009;7(2):355-359.

64. Burton RA, Tsurupa G, Hantgan RR, Tjandra N, Medved L. NMR solution structure, stability, and interaction of the recombinant bovine fibrinogen alphaC-domain fragment. *Biochemistry*. Jul 2007;46(29):8550-8560.

65. Hatzenbuehler J, Pulling TJ. Diagnosis and management of osteomyelitis. *Am Fam Physician*. Nov 2011;84(9):1027-1033.

66. Rydén C, Maxe I, Franzén A, Ljungh A, Heinegård D, Rubin K. Selective binding of bone matrix sialoprotein to Staphylococcus aureus in osteomyelitis. *Lancet*. Aug 1987;2(8557):515.

67. Persson L, Johansson C, Rydén C. Antibodies to Staphylococcus aureus bone sialoprotein-binding protein indicate infectious osteomyelitis. *Clin Vaccine Immunol*. Jun 2009;16(6):949-952.

68. Zaia J, Boynton R, Heinegård D, Barry F. Posttranslational modifications to human bone sialoprotein determined by mass spectrometry. *Biochemistry*. Oct 2001;40(43):12983-12991.

69. Vincent K, Durrant MC. A structural and functional model for human bone sialoprotein. *J Mol Graph Model*. Feb 2013;39:108-117.

70. Baht GS, O'Young J, Borovina A, et al. Phosphorylation of Ser136 is critical for potent bone sialoprotein-mediated nucleation of hydroxyapatite crystals. *Biochem J*. Jun 2010;428(3):385-395.

71. Martinez MD, Schmid GJ, McKenzie JA, Ornitz DM, Silva MJ. Healing of non-displaced fractures produced by fatigue loading of the mouse ulna. *Bone*. Jun 2010;46(6):1604-1612.

72. Blom AM, Hallstrom T, Riesbeck K. Complement evasion strategies of pathogens - Acquisition of inhibitors and beyond. *Molecular Immunology*. 2009; 46:2808-2817

73. Edwards AO, Ritter R, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science*. Apr 2005;308(5720):421-424.

74. Richards A, Kavanagh D, Atkinson JP. Inherited complement regulatory protein deficiency predisposes to human disease in acute injury and chronic inflammatory statesthe examples of vascular damage in atypical hemolytic uremic syndrome and debris accumulation in age-related macular degeneration. *Adv Immunol*. 2007;96:141-177.

75. Perkins SJ, Nan R, Li K, Khan S, Miller A. Complement factor H-ligand interactions: self-association, multivalency and dissociation constants. *Immunobiology*. Feb 2012;217(2):281-297.

76. Schmidt CW. FDA proposes to ban cephalosporins from livestock feed. *Environ Health Perspect*. Mar 2012;120(3):A106.